IJCSIS Vol. 12 No. 7, July 2014
ISSN 1947-5500

International Journal of Computer Science & Information Security


Cornell University
Library


Cogprints


Google scholar


.docstoc
find and share professional documents


ScientificCommons

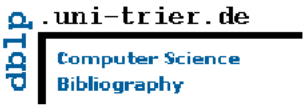

View my documents on
Scribd


BASE
Bielefeld Academic Search Engine


SCIRUS
search engine for science


SciRate.com


CiteSeerx beta


dblp .uni-trier.de
Computer Science
Bibliography


Q·Sensei


DOAJ
DIRECTORY OF
OPEN ACCESS
JOURNALS


EBSCO
HOST


ProQuest

IJCSIS

ISSN (online): 1947-5500

Please consider to contribute to and/or forward to the appropriate groups the following opportunity to submit and publish original scientific results.

CALL FOR PAPERS
International Journal of Computer Science and Information Security (IJCSIS)
January-December 2014 Issues

The topics suggested by this issue can be discussed in term of concepts, surveys, state of the art, research, standards, implementations, running experiments, applications, and industrial case studies. Authors are invited to submit complete unpublished papers, which are not under review in any other conference or journal in the following, but not limited to, topic areas.
See authors guide for manuscript preparation and submission guidelines.

Indexed by Google Scholar, DBLP, CiteSeerX, Directory for Open Access Journal (DOAJ), Bielefeld Academic Search Engine (BASE), SCIRUS, Scopus Database, Cornell University Library, ScientificCommons, ProQuest, EBSCO and more.

Deadline: see web site
Notification: see web site
Revision: see web site
Publication: see web site

Context-aware systems
Networking technologies
Security in network, systems, and applications
Evolutionary computation
Industrial systems
Evolutionary computation
Autonomic and autonomous systems
Bio-technologies
Knowledge data systems
Mobile and distance education
Intelligent techniques, logics and systems
Knowledge processing
Information technologies
Internet and web technologies
Digital information processing
Cognitive science and knowledge
Agent-based systems
Mobility and multimedia systems
Systems performance
Networking and telecommunications
Software development and deployment
Knowledge virtualization
Systems and networks on the chip
Knowledge for global defense
Information Systems [IS]
IPv6 Today - Technology and deployment
Modeling
Software Engineering
Optimization
Complexity
Natural Language Processing
Speech Synthesis
Data Mining

For more topics, please see web site https://sites.google.com/site/ijcsis/

For more information, please visit the journal website (https://sites.google.com/site/ijcsis/)

Editorial
Message from Managing Editor

*The **International Journal of Computer Science and Information Security** (IJCSIS) presents research, review and survey papers which offer a significant contribution to the computer science knowledge, and which are of high interest to a wide academic/research/practitioner audience. Coverage extends to all main-stream and state of the art branches of computer science, security and related information technology applications. As a scholarly open access peer-reviewed journal, IJCSIS mission is to provide an outlet for quality research publications. It aims to promote universal access with equal opportunities for international scientific community; to scientific knowledge, and the creation, and dissemination of scientific and technical information.*

*IJCSIS archives all publications in major academic/scientific databases. Indexed by the following International agencies and institutions: Google Scholar, Bielefeld Academic Search Engine (BASE), CiteSeerX, SCIRUS, Cornell's University Library EI, Scopus, DBLP, DOI, ProQuest, EBSCO. Google Scholar reported increased in number cited papers published in IJCSIS (**No. of Cited Papers:520, No. of Citations:981, Years:5**). Abstracting/indexing, editorial board and other important information are available online on homepage. This journal supports the Open Access policy of distribution of published manuscripts, ensuring "free availability on the public Internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of [published] articles".*

*IJCSIS editorial board, consisting of international experts, ensures a rigorous peer-reviewing process. We look forward to your collaboration. For further questions please do not hesitate to contact us at **ijcsiseditor@gmail.com**.*

,

A complete list of journals can be found at:
http://sites.google.com/site/ijcsis/

IJCSIS Vol. 12, No. 7, July 2014 Edition

ISSN 1947-5500 © IJCSIS, USA.

Journal Indexed by (among others):

IJCSIS EDITORIAL BOARD

TABLE OF CONTENTS

3. Paper 30061404: A Self-Training with Multiple CPUs Algorithm for Load Balancing using Time estimation (pp. 17-21)
Full Text: PDF

Aziz Alotaibi, Fahad Alswaina
Department of Computer Science, 221 University Ave, University of Bridgeport, Bridgeport, CT, USA

Abstract - In this paper, we propose a self-trading algorithm using two new parameters: time execution and type of priority to improve the load balancing performance. Load balancing uses information such as CPU load, memory usage, and network traffic which has been extracted from previous execution to increase the resource's utilization. We have included time execution for each property individually such as CPU bound, and Memory bound to balance the work between nodes. Type of priority has been taken into account to enhance and expedite the processing of request with high priority.

Keywords – Cloud Computing, Load Balancing, Resource allocation.

4. Paper 30061405: Result-Oriented Approach for Websites Accessibility Evaluation (pp. 22-30)
Full Text: PDF

Marya Butt, School of Governance, Utrecht University, Utrecht, the Netherlands

Abstract — The paper attempts to devise a result oriented approach for evaluating the accessibility of three Dutch government websites. Most of the research work pertaining website accessibility evaluation is intended to benchmark the organizations, however this study plans to initiate learning for the selected Government Bodies (GB) to improve websites accessibility. The devised approach spans three phases and is tested in three government bodies of the Netherlands. In the first phase, websites accessibility is evaluated for the selected government bodies. In the second phase, feedback from the web developers of the selected government bodies is collected to disclose their knowledge and practices. The third phase accentuates on measuring the results utilization. The websites evaluation is carried out according to the WCAG version 2.0 (level AA) by using various online tools - e.g. TAW, CCA (Color Contrast Analyzer), RIC (Readability Index Calculator) - and a test case to check that website is keyboard operable. Test results show that the selected websites failed to adhere to the WCAG 2.0. The feedback of the web developers revealed that though they are aware of these guidelines, yet clients do not want to compromise on other aspects, e.g. outlook and cost. The study initiated learning for all tested government bodies. Government bodies found the accessibility reports useful and showed perseverance to exploit research results in improving website accessibility.

Keywords-component; E-government, Websites Accessibility, Evaluation, Netherlands, WCAG 2.0

5. Paper 30061411: Performance Evaluation of Forward Difference Scheme on Huffman Algorithm to Compress and Decompress Data (pp. 31-36)
Full Text: PDF

Adamu Garba Mubi, Computer Science Department, Federal Polytechnic, Mubi, Adamawa State, Nigeria
Dr. P. B. Zirra, Computer Science Department, Federal University Kashere, Gombe State, Nigeria

Abstract - Data Compression using Forward Difference Techniques on Huffman algorithm is a research work which investigated how Forward Difference Techniques was used on Huffman to compress and decompress data without loss of information. The study measured the performance of Huffman algorithm against the Forward Difference on Huffman using Compression Ratio, Compression Factor and Saving Percentage. During the encoding the new algorithm reads the input file, serializes the distinct characters, determines the probability of each character, computes Forward Difference on the positions of each character, computes twos complement on the resulting difference, computes the new probability using the twos complement code, determines the codeword for each distinct character and finally determine the binary symbols to be transmitted. While decoding the new algorithm reads the whole encoded message bit-by-bit, determines a codeword from the coded message and determines a symbol the codeword represented; using the new probability the twos complement code is regenerated. Decimal

equivalent of the twos complement described a delta difference. Backward difference is used to determine the character positions of each character which is used again to reconstruct the whole message file. The results obtained revealed clearly that the performance of Forward Difference on Huffman is better than that of Huffman alone.

Keyword: Data, Huffman algorithm, data compression

6. Paper 30061422: Wireless Sensor Networks Attacks and Solutions (pp. 37-40)
Full Text: PDF

Naser Alajmi, Computer Science and Engineering Department, University of Bridgeport, Bridgeport, CT 06604, USA

Abstract — A few years ago, wireless sensor networks (WSNs) used by only military. Now, we have seen many of organizations use WSNs for some purposes such as weather, pollution, traffic control, and healthcare. Security is becoming on these days a major concern for wireless sensor network. In this paper I focus on the security types of attacks and their detection. This paper anatomizes the security requirements and security attacks in wireless sensor networks. Also, indicate to the benchmarks for the security in WSNs

Keywords-Wireless sensor network, security, vulnerability, attacks

7. Paper 30061423: Enhancing the Accuracy of Biometric Feature Extraction Fusion Using Gabor Filter and Mahalanobis Distance Algorithm (pp. 41-48)
Full Text: PDF

Ayodeji S. Makinde, Yaw Nkansah-Gyekye, Loserian S. Laizer
School of Computational and Communication Science and Engineering, NM-AIST, Tanzania

Abstract - Biometric recognition systems have advanced significantly in the last decade and their use in specific applications will increase in the near future. The ability to conduct meaningful comparisons and assessments will be crucial to successful deployment and increasing biometric adoption. The best modality used as unimodal biometric systems are unable to fully address the problem of higher recognition rate. Multimodal biometric systems are able to mitigate some of the limitations encountered in unimodal biometric systems, such as non-universality, distinctiveness, non-acceptability, noisy sensor data, spoof attacks, and performance.
More reliable recognition accuracy and performance are achievable as different modalities were being combined together and different algorithms or techniques were being used. The work presented in this paper focuses on a bimodal biometric system using face and fingerprint. An image enhancement technique (histogram equalization) is used to enhance the face and fingerprint images. Salient features of the face and fingerprint were extracted using the Gabor filter technique. A dimensionality reduction technique was carried out on both images extracted features using a principal component analysis technique. A feature level fusion algorithm (Mahalanobis distance technique) is used to combine each unimodal feature together. The performance of the proposed approach is validated and is effective.

Keywords – Gabor filters; Mahalanobis distance; principal component analysis; face; fingerprint; feature extraction.

8. Paper 30061428: GPGPU based Parallel Spam Filter (pp. 49-58)
Full Text: PDF

Prachi Goyal Juneja, M.Tech Scholar, Maulana Azad National Institute of Technology Bhopal (M.P) India-462003
R. K. Pateriya, Associate Professor, Maulana Azad National Institute of Technology Bhopal (M.P) India-462003

Abstract - Spam means unwanted emails in our mailboxes each day. These emails consist of promotional messages from companies, viruses, lucrative offers of earning extra income and many more. They are sent in bulk to flood our

mailboxes and come from unknown sources. Various ways have been devised to deal with spam; these are known as Spam Filtering Techniques. Spam Filtering is done based on many parameters like keywords, URL, content etc. Content based spam filtering is becoming famous since it incorporates the judging of the email content and then analyzing it to be spam or ham. As the data is increasing and electronic data taking over most of the communication medium, one needs faster processing and computing devices. GPGPU's have come up in a great way in sharing the CPU's tasks and make parallel processing possible.

Keywords- Spam, Bayesian Spam Filtering, Serial Spam Filter, Parallel Spam Filter, Spamicity.

9. Paper 30061429: Mobile- Health Application Software Design and Development (pp. 59-66)
Full Text: PDF

Ayangbekun Oluwafemi J., Department of Information Systems, University of Capetown, South Africa
Kasali Olanrewaju M., Department of Information Technology, Crescent University Abeokuta, Nigeria

Abstract — Mobile technologies are fast developing and it has completely changed the way we interact and provide healthcare services. The rapid spread of mobile technologies and inventive applications to address health related problems has evolved into a new field known as mobile-Health. The purpose of this research is to improve the quality and access to health care services with the aid of mobile-Health application software known as "Crescent Mobile Health". This paper will address the problem of self medication by creating a channel of communication between a patient and doctor at distant environment there by solving emergency situations. The method used to address this problem is by designing and developing mobile-Health application software, which can be used by patients via an android smartphone that is used to communicate with a doctor/pharmacist/laboratory scientist using electronic-Health application software known as Crescent Health Information System on a desktop via the intranet. The two applications on smartphone and desktop are able to communicate via instant messaging by a persistent connection known as "sockets" and "pusher" which provides implementation for interconnectivity. The Crescent Health Information System can carry out major functionalities such as drugs and tests inventory, instant messaging, prescriptions of drugs, prescription of tests and profile update. The Crescent Mobile Health can also carry out functionalities such as instant messaging, viewing of prescribed drugs, tests, health tips and help file. The mobile-Health application software was developed using java programming language and android development studio while the electronic-Health (E-Health) application software was developed using PHP programming language and MYSQL database. The results of the development of this project concludes that mobile-Health application software has been able to resolve the problem of communication between a patient and a doctor and has provided a means to verify drugs available and tests carried out in the clinic/health sector.

Keywords - Electronic-Health; Healthcare; Intranet; Mobile- Health; Patient; Smartphone; Socket

10. Paper 30061427: System Analysis and Design for integrated sponsored SMS/USSD Based M-Services: A case study of Maternal Health M-Service in Tanzania (pp. 67-77)
Full Text: PDF

Timothy Y. Wikedzi, Ramadhani S. Sinde
Computational and Communication Sci & Eng, Nelson Mandela African Institution of Sci & Tech, Arusha, Tanzania
Dan K. McIntyre, Information Technology, University of Iringa, Iringa, Tanzania

Abstract -- Mobile phones have proven to be the best way of providing reliable access to information to people in low and mid income countries where other forms of communication perform poorly. As a result of the wide spread of mobile phones, there has been an increase in number of Mobile Application (M-Services) which are being used as a tool for disseminating different type information to people. M-Services of this nature are established to address informational challenges that are faced by people especially low income people. Because of this then, these projects must be sustained so that people can enjoy the benefits of it. Contrary to this, reports show that most of these M-Services are facing the challenge of cost of operating them, which in a direct way affects the sustainability of these services. In this paper therefore we present an analysis and later design of a noncommercial M-Service, which integrates advertising functionality as a tool for subsidizing the cost of operating M-Services. To achieve this we

have employed some concepts of Information System Analysis and Design (ISAD) as the guiding principle towards achieving our design. A prototype of M-Health is used for the study.

Keywords-M-Service; ISAD; Ad, USSD; SMS; Mobile; Sustainable; Cost of operation.

11. Paper 30061407: Conversion of an SR-Flip Flop to a JK-Flip Flop (pp. 78-92)
Full Text: PDF

Prof. Olawale J. Omotosho, Engr. Samson O. Ogunlere
Babcock University, Computer Science Department, Ilishan-Remo, Ogun State, Nigeria

Abstract - This paper presents a design method to convert a conventional SR-Flip Flop to perform the functions of a corresponding conventional JK-Flip Flop. This requirement becomes very necessary because of the many applications of JK-Flip Flops in digital systems, especially in those systems that drive production industries. In such industries, uninterrupted production is one of the targets required to pay attention to in order not to lose production and consequently revenue. Equipment failure can be responsible for such an unwanted state of production. Therefore, downtime of any equipment becomes very crucial in the assurance procedure of associated equipment and instrumentation of a manufacturing plant. The cause of a large downtime of any equipment is mainly due to unavailability of spare parts and sometimes incompetence and inexperience of the Technologists responsible for the up-keep and assurance of these equipment and instrumentation. Technologist must be versatile in providing alternative solutions to existing provisions that is adequate to solve any prevailing situation which requires urgent attention to keep production going. Such experience is not only borne out of hands-on practice but can be acquired by sound theoretical knowledge of what to do. This paper examines a situation where a device (JK-Flip Flop) is not available to replace a defective one but an SR-Flip flop is configured to be used for the same purpose without degradation of performance.

Keywords—Conventional Flip-Flops, uninterrupted production, unavailability of spare parts, incompetence and inexperience of the Technologists, downtime of equipment, K-maps.

12. Paper 30061418: Digital Shorthand Based Text Compression (pp. 93-97)
Full Text: PDF

Yogesh Rathore, CSE,UIT, RGPV, Bhopal, M.P., India
Dr. Rajeev Pandey, CSE,UIT, RGPV, Bhopal, M.P., India
Manish K. Ahirwar, CSE,UIT, RGPV, Bhopal, M.P., India

Abstract — With the growing demand for text transmission and storage as a result of advent of net technology, text compression has gained its own momentum. Usually text is coded in yank traditional Code for data Interchange format. Huffman secret writing or the other run length secret writing techniques compresses the plain text[6][11]. We have planned a brand new technique for plain text compression, that is especially inspired by the ideas of Pitman Shorthand. In these technique we propose a stronger coding strategy, which can provide higher compression ratios and higher security towards all possible ways in which of attacks while transmission. The target of this method is to develop a stronger transformation yielding larger compression and additional security[11]. The basic idea of compression is to transform text in to some intermediate form, which may be compressed with higher efficiency and more secure encoding, that exploits the natural redundancy of the language in creating this transformation.

Keywords - Compression; Encoding; REL; RLL; Huffman; LZ; LZW; Pitman Shorthand; Compression;

Logical Analysis of an Accelerated Secure Multicast Authentication Protocol

Ghada Elkabbany
Informatics Dept., Electronics Research Institute
Cairo, Egypt

Mohamed Rasslan
Informatics Dept., Electronics Research Institute
Cairo, Egypt

Heba Aslan
Informatics Dept., Electronics Research Institute
Cairo, Egypt

Abstract— **Multicast authentication is a challenging problem, because it should verify the received packets without assuming the availability of the entire original stream and resist many types of attacks, such as pollution attacks. Researchers have proposed many solutions in literature with major drawbacks in high communication and computation overheads. Others suffer from packet loss and pollution attacks. Recently, signature techniques were used to provide multicast authentication. Signcryption techniques have the advantage of achieving the basic goals of encryption and signature schemes. But, it suffers from the inability to resist packet loss. In a previous work, we proposed a multicast authentication protocol that is based on signcryption techniques and erasure code function to solve the packet loss problem. In this paper, we utilize pipelining technique to reduce the computation overhead. Pipelined technique is chosen due to its suitability for signcryption algorithm nature. The pipelined technique reduces the computation time. Moreover, a verification of our protocol using BAN logic is performed. The analysis shows that it achieves the goals of authentication without bugs or redundancies. A comparison of multicast authentication protocols is carried out. The results show that the accelerated multicast authentication protocol resists packet loss and pollution attacks with low computation and communication overheads, therefore, it could be used in real-time applications.**

Keywords- Multicast Communication, Authentication, Signcryption, Erasure Code Functions, Parallel Pipelined, BAN Logic.

I. INTRODUCTION

Multicast communication accomplishes efficient exchange of messages. However, multicast protocols do not provide any mechanisms for providing confidentiality or authenticity of the received messages. Applications that take advantage of multicast communication include: video conferencing, distance learning, corporate communications, and distribution of software, stock quotes and news. The multicast authentication is a serious problem. Authenticity means that the recipient could verify the identity of the sender and ensures that the received message comes from the supposed originator. For multicast communication, authentication is a challenging problem, since it requires the verification of data originator by a large number of recipients. Multicast authentication protocols must have the following characteristics: it must defend against both packet loss and pollution attacks. In addition, it must have low computation and communication overheads. Researchers have proposed many solutions in literature. The major drawback of some of these solutions was the high communication and computation overheads. Others suffer from packet loss and pollution attacks. Recently, signature techniques were used to provide multicast authentication. Signcryption techniques have the advantage of achieving the basic goals of encryption and signature schemes such as: confidentiality, authenticity, and non-repudiation. On the other hand, it suffers from the inability to resist packet loss. Since, to perform authentication, all the packets must be delivered to the designated recipients.

In a previous work, we proposed a multicast authentication protocol [1], which is based on signcryption techniques that solve the problem of pollution attacks and lower the communication overhead. This protocol uses erasure code function to solve the packet loss problem. In this paper, in order to accelerate our proposed protocol, we use a pipelined technique that reduces the computation overhead of the proposed signcryption algorithm. Pipelined concept which is chosen due to its suitability for signcryption algorithms reduces the computation time with respect to its corresponding values of a sequential execution. Moreover, a verification of our proposed protocol using BAN logic is performed. Finally, a comparison of multicast authentication protocols is presented. The comparison is undertaken for two cases: first, without parallelization, while, in the second case, parallelization is incorporated to enhance the computation overhead. This paper is organized as follows: in the next section, background and related work are detailed. Then, a description of our proposed multicast authentication protocol, the pipelined technique, and finally, the verification using BAN logic are given in Section 3. In Section 4, a comparison of the proposed protocol with other protocols is discussed. Finally, Section 5 concludes the paper.

II. RELATED WORK

A. Multicast Authentication

Unlike broadcast network addressing method that delivers information to all the members in a network, information in multicast network addressing method is sent to a set of chosen destinations only at once. In Internet, which is

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

an open network, session key distribution is the primary step to accomplish multicast security services. Matsuura *et al.* [2] proposed a scalable and compact multicast key distribution protocol using signcryption to obtain a secure and authenticated key delivery service. For a sparsely distributed network, the most scalable techniques use a shared-tree approach, such as the core-based tree (CBT) routing protocol [3]. Shared-tree technique is scalable and has fewer entries in routing tables. To achieve a scalable routing scheme, Matsuura *et al.* made use of the core-based tree (CBT) [3] routing protocol and multiple key distribution centers (KDCs). These techniques are more practical and flexible than a single trusted key distribution center. Matsuura *et al.* protocol in [2] got benefits of signcryption to significantly reduce the communication cost and computation time delay of the CBT protocol. In Matsuura *et al*. protocol, a trusted third party, key distribution center (KDC), aims to issue a key to valid users in the network. To avoid a single point of failure, a group key distribution center (GKDC) has been proposed as a practical solution for key distribution over the Internet. On the other hand, distribution environment gives a chance to improve the performance by applying parallelism technique to signcryption protocols. In GKDC, each local group KDC has its own public and private signing key pair. Some look at CBT as a routing protocol that employs GKDCs. Before accessing the services that legitimate users wants to get across groups, legitimate users have to join the new group and receive a distributed key. CBT conventional routing protocol uses cascaded signature and public-key encryption schemes to achieve the origin authenticity and secrecy of transferred packets. Instead of using encryption/signing process, signcryption can be applied in place of separate encryption and signing to reduce both communication bandwidth and computational time overheads. Group Key Distribution is used in multicast communications. The authentication of multicast communication is a challenging problem. Any authentication scheme for multicast streams should verify the received packets without assuming the availability of the entire original stream. In addition, it should resist against many types of attacks by an adversary, such as pollution attacks. Researchers have proposed many solutions in literature.

There are two approaches to solve the multicast authentication problem. The first is to design more efficient signature schemes and amortize the cost of signature over several packets. In this approach, efficient digital signature schemes have been proposed in [4, 5]. These schemes are impractical for real-time applications due to the fact that they suffer from the communication overhead problem. We have to mention that, these schemes overcome the computational problem. The second approach is to amortize signature over several packets as proposed in [5-9]. Gennaro and Rohatgi [6] did an early work to amortize signature over several packets. In their work, the packets stream is divided into blocks of "*L"* packets. Moreover, Gennaro and Rohatgi use a chain of hashes to link each packet to the predecessor one and the last hash is signed. This approach has the advantage of solving the computation and communication overheads problem. But, it assumes no loss of any packet during the transmission. This is an impractical assumption and cannot be achieved in real life. Almost all multicast applications use Internet Protocol (IP) addressing. IP is a UDP and no guarantee to prevent packet loss. This raises the need for a technique to prevent packet loss in multicast authentication protocols that amortize signature over several packets. Golle and Modadugu [10] proposed a solution for the packet loss problem by appending the hash of each packet into two different places. The first copy of the hash is located in the next packet. The second is located in the packet that succeeds the current packet by *a* places and only the final packet is signed. Golle and Modadugu's idea [10] is valid due to the fact that packet loss over the Internet occurs in bursts behavior [7]. Also, Golle and Modadugu's idea [10] can resist several bursts of a certain number of packets. Golle and Modadugu [10] do not illustrate how the packet that contains the signature is sent.

The signature-packet loss requires its retransmission several times by the sender. When different receivers lose different sets of packets, senders retransmit these packets. This retransmission could overload the resources of both the sender and the network. To solve the problem of packet loss, Wong and Lam [5] proposed another solution. Wong and Lam [5] divide the packet stream into blocks of *"L"* packets and construct a tree of hashes of degree 2. In [5], Wong and Lam calculate the hash values of the *"L"* packet. These hash values correspond to the leaves of the tree and only the root of the tree is signed by the sender. The idea behind the Wong and Lam [5] is that each packet carries the information required for its authentication. The packet signature and the siblings of each node along its path to the root are appended to the corresponding packet. The receiver verifies the signature to authenticate the corresponding packet. If a packet arrived after a loss of its predecessor and due to the fact that each packet carries the required information for its authentication process, lost packet will not affect the ability of the receiver to authenticate the arrived packet. This solution suffers from a high communication overhead because it requires the sender to append $log_2(L)+1$ hashes and the signature to each packet. Perrig *et al.* [8, 9] proposed efficient solutions for the authentication problem that are based on MACs and revealing the MAC keys after a certain time interval. Perrig *et al.*'s solutions have low computation and communication overheads. But, their solutions have two drawbacks. The first is that requires the sender and the receivers maintain the synchronization of their clocks. The second is that suffer from multiple sent of signature packet in case of packet loss.

Using erasure codes, Pannetrat and Molva [11] proposed a solution to the problem of sending multiple signature packets and packet loss. Given that the loss rate does not exceed a certain limit, erasure codes [12, 13] allow the receiver to restore the original data. The sender divide the stream into blocks each consists of *L* packets. Also, the sender calculates the hashes of these packets. Then, the output hash values are input to an erasure code function to produce X,

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

which consists of the hash values and extra information (E) in order to resist transmission loss. For these packet hashes, a signature (S) is computed. Both E and S are input to the second erasure code function. The output of the second erasure code function is divided into *L* fragments. These fragments are appended to each corresponding packet. Erasure codes manipulation enables the recipient to restore both the hash values and the signature in case of loss. In [14], a similar solution is presented and named Signature Amortization using Information Dispersal Algorithm (SAIDA). Park *et al.* [14] use erasure code as one stage that contains the packet hashes and the signature as input. Erasure codes can resist only the packet loss threat model, where packets are assumed to be lost but not corrupted in transit. But, packets could be modified, delayed, dropped, and lost. In, [15], these threats are defined as pollution attacks. Karlof *et al.* [15] propose a solution to pollution attacks and is known as Pollution Resistant Authenticated Block Streams (PRABS). Karlof *et al.*'s solution is based on SAIDA. In this solution, each output symbol of the erasure code is augmented by additional information. Karlof *et al.* call this additional information a witness information. Karlof *et al.*'s use this information to differentiate between legitimate symbols and invalid symbols. To generate witness information, Merkle hash tree is constructed. In Merkle hash tree, output symbols from the erasure code are considered as leaves of the tree. Each symbol is augmented by the siblings along its path to the root. This information is used to partition symbols as valid or invalid. Only valid symbols are used to restore the original packet hashes and the corresponding signature. Karlof *et al.* solution overcomes the pollution attack problem. But, it has a large communication overhead compared to the abovementioned multicast authentication protocols.

Besides, in public key cryptography, encryption and signature schemes are fundamental primitives that provide privacy and authenticity. Cryptographers used to consider these two primitives as discrete building blocks that have to be designed and studied independently. On the contrary, there are numerous settings where both primitives are required to accomplish both privacy and authenticity by comprising the known solutions of each of the two components. The sign-then-encrypt and encrypt-then-sign paradigms are main methods to implement authenticated encryption schemes [16, 17]. These schemes achieve the non-repudiation property but they are costly in terms of communication and computation. The term signcryption was originally presented and analyzed by Zheng in [18] with the primary goal of reaching greater efficiency than can be accomplished when performing the signature and encryption operations separately. Signcryption schemes [19, 20] aim to simultaneously accomplish the primary goals of encryption and signature schemes, namely confidentiality and authenticity. Moreover, signcryption schemes must accomplish non-repudiation, which assures that the sender of a message cannot later deny that she has sent the message. Namely, the recipient of a message can persuade a third party that the sender definitely sent the message. In addition to sustaining the confidentiality, authenticity and non-repudiation properties, some signcryption schemes are designed to accomplish forward secrecy and past recovery [18]. The deployment of signcryption in multicast authentication is categorized into two approaches [21]. The first solution necessitates the signcryption of a message for *n* receivers. While this solution lowers the computation overhead, it suffers from the large communication overhead and the lack of resistance to packet loss. An example of this solution could be found in [22, 23]. The second solution is to deploy the technique of randomness [24]. Again this solution could not defend packet loss because of the necessity to receive all packets to guarantee authentication. Duan *et al.* [21] suggested a scheme that is based on signcryption scheme which only requires one pairing computation to signcrypt a message for *n* receivers. This scheme lowers the communication and computation overheads without solving the problem of packet loss. In [22], Pang *et al.* suggested an anonymous multi-recipient signcryption scheme, which is based on bilinear operation. Although, this scheme uses one bilinear operation to signcrypt a message for *n* receivers, it suffers from the fact that it requires high computation and communication overheads. All the above mentioned signcryption schemes could not defend packet loss, since it necessitates the receiving of the whole block of packets in order to be able to ensure the authenticity of the received message. In the next section, different techniques used to improve of signcryption protocols by parallelization are illustrated.

B. Parallel and Pipelined Signcryption

The signcryption scheme requires one computation for "encryption" and one inverse computation for "authentication", which is of great practical significance in directly performing long messages, since the major bottleneck for many public encryption schemes is the excessive computational overhead of performing these two operations [25]. Hu et. al., in their work [25] presented a highly practical parallel signcryption scheme named PLSC from trapdoor permutations (TDPs for short) was built to perform long messages directly. This scheme followed the idea "scramble all, and encrypt small", using some *scrambling operation* on message *'L'* along with the user's identities, and then passing, in parallel, small parts of the scrambling result through

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

corresponding TDPs. This design enabled the scheme to flexibly perform long messages of arbitrary length while avoid repeatedly invoking TDP operations such as the CBC mode, or verbosely black-box composing symmetric encryption and signcryption, resulting in noticeable practical savings in both message bandwidth and efficiency. Cutting out the verbosely repeated padding, the newly proposed scheme is more efficient than a black-box hybrid scheme. This scheme has been proven to be tightly semantically secure under adaptive chosen ciphertext attacks and to provide integrity of ciphertext as well as non-repudiation in the random oracle model.

Privacy and authenticity are two basic security goals; there are many applications that require both goals to be achieved simultaneously. However, the main problem considered initially was the design of encryption and signature so that their concatenation maximizes savings of computing resources. The goal of parallel signcryption algorithms is to achieve the lower bound in terms of time necessary to perform authenticated encryption and decryption. The parallel encryption was introduced by An *et al.* [26], in their work they developed a security model for parallel signcryption and present the commit-then-encrypt-and-sign (*CtE*&*S*) scheme. This scheme used three cryptographic blocks: a commitment scheme, a public key encryption scheme, and a signature scheme. Another solution was given by Pieprzyk and Pointcheval [27]; this model implemented the commitment part very efficiently using secret sharing. It also showed how to combine encryption and signing so that they strengthen each other and can be executed in parallel. Han *et al.*, [28], proposed a parallel multi-recipient signcryption scheme. This scheme was called ParaSC-BLS and it is semantic secure in random oracle model assuming that the GDH problem is hard. With pre-computing, the speedup of ParaSC-BLS was up to N. As the number of recipients growing, the computational overheads of ParaSC-BLS were constant. Moreover, for the randomness reusing, ciphertext is aggregated and the total transmission overheads are reduced. A modified parallel multi-recipient scheme was proposed [29], this paradigm called ParaSC-GDH. It enhanced the performance when a sender sends distinct messages to multiple recipients in imbalanced wireless networks. Randomness reusing and ciphertext aggregation were used to reduce the computation and transmission overheads on the sender node. Their experiment result showed that, the improvement of CPU time and ciphertext sizes are 48.6% and 42.8% for 1,000 distinct messages. Ring signcryption is a cryptographic primitive which combines the functions of a ring signature scheme and an encryption scheme. It can provide confidentiality, anonymity and authenticity simultaneously. Zhu *et al.*, [30] presented a security flaw in a previous certificateless ring signcryption scheme (Wang *et al.*'s scheme) and proposed an efficient parallel ring signcryption scheme in certificateless public key setting. Compared with Wang *et al.*'s scheme, this scheme needed three less pairings and enjoyed a shorter ciphertext. Furthermore the parallel execution of the expensive signature and encryption operations implied a gain in running efficiency. In addition, this scheme is not only indistinguishable against chosen-ciphertext attacks but also existentially unforgeable against chosen-message attacks.
In our work [31], we used pipelined concept to speed-up the signcryption algorithm described in [32]. This leads to the decrease of the total computation time compared to its corresponding values of sequential execution. The proposed technique uses *'M''* processors to perform the signcryption operation. Then, the output of each function unit was shifted from a stage to the next stage. In the next section, a description of our multicast authentication protocol [1] is detailed. In order to lower the computation overhead, a pipelined technique [31] is implemented. Then, a logical analysis of the multicast authentication protocol is undertaken.

III. LOGICAL ANALYSIS OF ACCELERATED MULTICAST AUTHENTICATION PROTOCOL

A. The Multicast Authentication Protocol

Signcryption techniques are intended to simultaneously accomplish confidentiality, authentication and non-repudiation. Although, the use of signcryption techniques improves communication and computation overheads, this solution suffers from the incapability to overcome packet loss problem. This is due to the fact that all the block of packets must be received by the chosen recipients before performing authentication. In order to overcome the packet loss problem, the proposed protocol uses erasure code functions which allow the receiver to repair the original data under the condition that the loss rate does not exceed a firm value (*R*). In our work [1] we suggest a protocol that is based on the idea of amortizing the signature over the signcrypted text. This protocol uses signcryption technique to offer both confidentiality and authenticity, and to overcome pollution attacks. The signcryption protocol used in our solution is elaborated in [1, 32]. The scheme in [32] is more effective than all the previously presented schemes. It allows the recipient (verifier) to restore the message blocks upon receiving their corresponding signature blocks. The scheme is perfect for some application requirements and it fits packet switched networks.

Our protocol [1] assumes: the System Authority (SA) selects a large prime number *p* such that *p*-1 has a large prime factor *q*. In addition, SA picks an integer, *g*, with order *q* in *GF(p)*, and Lets "*f* " be a secure one way hash function. SA publishes *p*, *q*, *g* and *f*. Each user, U_i, chooses a secret key $x_i \in Z_q$ and computes the corresponding public key $y_i = g^{x_i} \bmod p$. In addition, all the group users share a secret key $x_{s\text{-}group}$ and its corresponding public key, $y_{p\text{-}group} = g^{x_{s-group}}$, which is used to encrypt group messages. The usage of a public group key will decrease both the communication and computation overheads, since it necessitates the message to be signcrypted once instead to be signcrypted *n* times (where *n* is the number of receivers). This pair of keys must be altered in case of any membership change, i.e. a new member joins/leaves the group. Examples of solutions to solve the key distribution problem in case of a member join/leave are presented in [33, 34]. When a sender *A* wants to send a

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

message to the whole group, it divides the stream into blocks of L packets ($Pack_1$, $Pack_2$, $Pack_3$,…, $Pack_{L-2}$, $Pack_{L-1}$, $Pack_L$). The value of these packets must be less than the value of p. The sender A, with secret key x_a and public key $y_a = g^{x_a}$, uses the following steps before sending the multicast message:

(1) Pick random numbers k, l Z_q^* and set $r_0 = 0$, then compute $y_{p-group}^{k}$ mod p and $t = g^k \bmod p$.
(2) Compute: $r_i = Pack_i \,.\, f\,(r_{i-1} \oplus y_{p-group}^{k}) \bmod p$, for $i = 1, 2, \ldots, L$.
(3) Compute: $s = k - r \,.\, x_a \bmod q$, where $r = f\,(r_1, r_2, r_3, \ldots, r_L)$.
(4) Then, the sender computes $c_1 = g^l \bmod p$ and $c_2 = r_L \,.\, y_{p-group}^{k} \bmod p$.
(5) Next, the sender applies the erasure code function on r, c_1, c_2, s and t. The output of the erasure code function is divided into L-1 parts ($T_1, T_2, \ldots, T_{L-1}$), where each part is appended to each packet output of the signcryption algorithm. Then, the sender broadcasts the following message: ($r_1 \parallel T_1$, $r_2 \parallel T_2$, .., $r_{L-1} \parallel T_{L-1}$). Figure 1 illustrates the steps required to perform the proposed protocol.

To reestablish and authenticate the received block, it is sufficient to receive $L(1\text{-}R)$ packet. Loss in one packet will only affect the recovery of this packet and the following one, the remainder of the block could be recovered. After receiving the sent message, each recipient checks the signature by comparing $t^{x_s - group}$ to $y_{p-group}^{s} \cdot y_{ap-group}^{r} \bmod p$, where $y_{ap\text{-}group} = y_a^{x-group} \bmod p$. If the check doesn't hold, this indicates that the received packets are modified and must be discarded. On the other hand, if the check holds, then each recipient calculates $r_L = c_2 \,.\, c_1^{-x_b} \bmod p$. Finally, each recipient restores message blocks using the following equation: $P_i = r_i \,.\, f\,(r_{i-1} \oplus t^{x_b})^{-1} \bmod p$, for $i = 1, 2, \ldots, L$ and $r_0 = 0$. The proposed protocol provides both confidentiality and authenticity in one step. Therefore, the computation overhead decreases, this makes the proposed protocol appropriate for real-time applications. To decrease the communication overhead, which is considered one of the major disadvantages of using signcryption techniques, we use a pair of group public key. This eliminates the need to encrypt the message using each recipient's public key and as a result, lowers the communication overhead. Other advantage of the proposed protocol is that it could resist both packet loss and pollution attacks with low computation and communication overheads. In the next section, a comparison of the proposed protocol with other multicast authentication protocols is presented. As mentioned in the previous section, parallelizing the computation of cryptographic algorithms is a promising approach to reduce the execution time and eventually the energy consumption of such algorithms. In this work, the proposed protocol [1, 32] has been accelerated using pipelining algorithm [31]. The results showed that the proposed pipelined technique reduced the computation time required to execute the selected signcryption algorithm by approximately 72%, compared to its corresponding values of sequential execution.

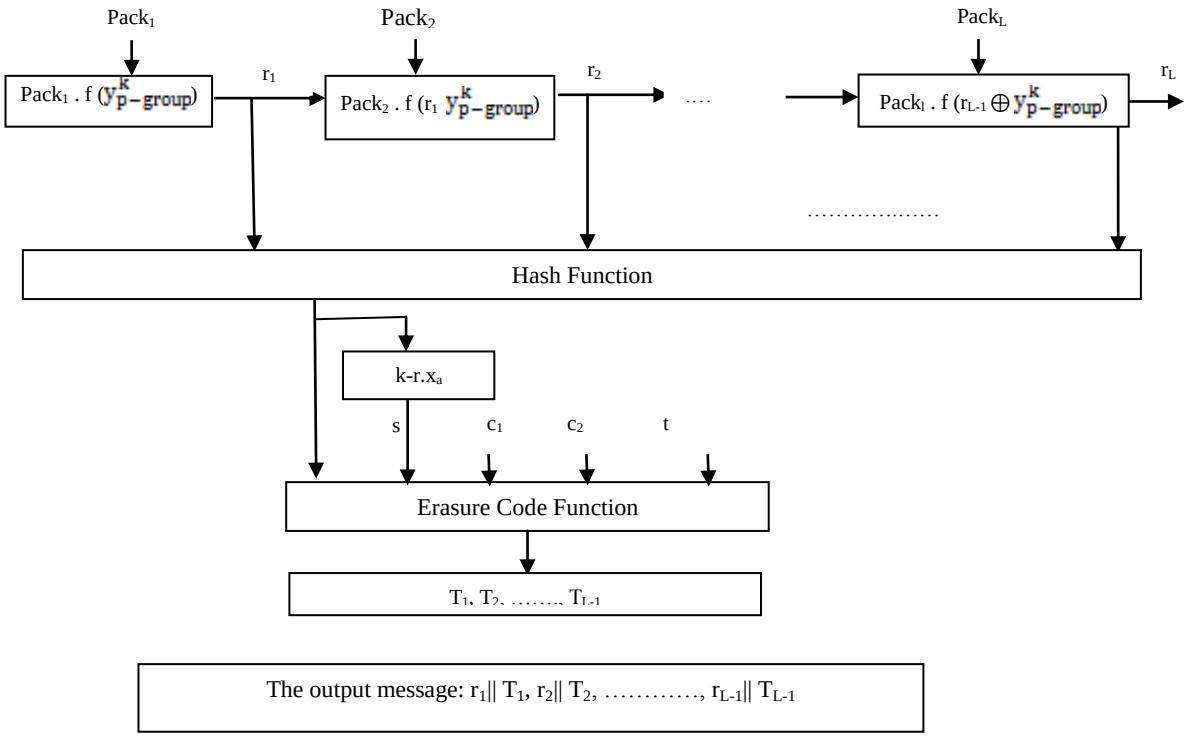


Figure 1. The multicast authentication protocol architecture.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

B. Verification of the Multicast Protocol using BAN Logic

Authentication protocols are the basis of security in many distributed systems, and it is therefore essential to ensure that these protocols function correctly. Unfortunately, their design has been extremely error prone. Most of the protocols found in the literature contain redundancies or security flaws [35]. In their work [35], M. Burrows et. al proposed a method that uses the logic to describe the authentication protocols. They transformed each message into a logical formula which is an idealized version of the original message. In this section, a logical analysis of the multicast protocol described in Section 3.1 using BAN logic will be presented. For a successful verification of the protocol, the belief state of communicating parties should satisfy the protocol goals. We will consider the multicast communication is completed between principals A and B, if there is a data packet "X" which the receiver B believes that it is sent by A. Thus, authentication between A and B will be completed if B $\mid\equiv$ A $\mid\equiv$ X, and B $\mid\equiv$ X, where the symbol $\mid\equiv$ means believes. First, the basic rules of the BAN logic are listed below:

- The interpretation rule

$$\frac{P \mid\equiv (Q \mid\sim (X,Y))}{P \mid\equiv (Q \mid\sim X), P \mid\equiv (Q \mid\sim Y)}$$

The above rule means that if P believes that Q once said a message containing both X and Y, therefore it believes that Q once said each statement separately.

- Message Meaning Rule

$$\frac{P \equiv \xrightarrow{k} Q, P \lhd [X]_{K^{-1}}}{P \mid\equiv Q \mid\sim X}, P \neq Q$$

This means that if P believes that K is the public key of Q, and P sees a message X signed by K^{-1}, this implies that P believes that Q once said X.

- Nonce Verification Rule

$$\frac{P \mid\equiv \#(X), P \mid\equiv Q \sim X}{P \mid\equiv Q \mid\equiv X}$$

The above rule means that if P believes that X is a recent message and Q once said X, therefore it believes that Q believes in X.

- Jurisdiction Rule

$$\frac{P \mid\equiv Q \Rightarrow X, P \mid\equiv Q \mid\equiv X}{P \mid\equiv X}$$

This rule means that if P believes that Q has jurisdiction over X, and P believes that Q believes in X, then P believes in X.

- Freshness Rule

$$\frac{P \mid\equiv \#(X)}{P \mid\equiv \#(X,Y)}$$

The above rule means that if P believes in the freshness of X and Y, therefore it believes in the freshness of each statement separately.

The analysis is undertaken for the message exchanged between the sender and one of the receivers. For the other group members, the same analysis could be carried out. The authentication is considered completed between A and B, if the following goals are achieved:

Goal 1: B $\mid\equiv$ A $\mid\equiv$ Pac

Goal 2: B $\mid\equiv$ Pac

Where, *Pac* represents the group of packets sent by A. In order to complete the analysis, the following assumptions are made:

$$B \mid\equiv \xrightarrow{k_a} A \quad (1)$$

$$B \mid\equiv A \Rightarrow Pac \quad (2)$$

$$B \mid\equiv \#t \quad (3)$$

Equation (1) indicates that B believes that K_a is the public key of A. Then, equation (2) indicates that both B believes that A has jurisdiction over the group of packets sent. Finally, equation (3) indicates that B believes in the freshness of t (since it is changed for each group of packets). After making the assumptions, the messages transferred in the initial phase are transformed into logical formulas. Finally, the basic rules of the BAN logic will be applied to the logical formulas. Following is the transformation of the proposed protocol into logical formulas:

$$A \longrightarrow B: \{\{Pac\}_{K_a^{-1}}, c_1, c_2, t\} \quad (4)$$

The analysis of the protocol can now be performed. By applying message meaning rule to equation (4) and using equation (1), the following can be deduced:

$$B \mid\equiv A \mid\sim (Pac,t)$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

But, B believes in the freshness of t (equation (3)). Thus, applying nonce verification rule, the following is obtained:

$$B \mid\equiv A \mid\equiv Pac \qquad (5)$$

Then, by applying jurisdiction rule using equation (2), the following is obtained:

$$B \mid\equiv Pac \qquad (6)$$

From equations (5) and (6), one can deduce that the proposed protocol achieves the goals of authentication without bugs or redundancies. In the next section, comparison of the proposed protocol with other multicast authentication protocols is detailed.

IV. COMPARISION OF PROPOSED PROTOCOL WITH OTHER MULTICAST AUTHENTICATION PROTOCOLS

In this section, a comparison between Wong-Lam, PRABS, Pang *et al.* and our protocol [1, 32] is presented. In order to perform the comparison the following general assumptions are considered:

- The stream to be authenticated is divided into blocks of *L* packets (of length smaller than that of the prime number *p* mentioned in the previous section).
- The calculations are specified for the authentication of one block.
- The loss rate must be less than *R*.
- The length of *p* = length of *q* = *Len_P*.
- The length of the hash functions output equals '*Hout*', the signature length equals '*Sigout*', and the length of the erasure code output '*Eout*' is given by:

  Eout = Ein * (1+R)

  where, *Ein* represents the length of the erasure code input and *R* corresponds to the loss rate.

The comparison will be undertaken according to the following criteria:

- The computation overhead: the processing needed at the sender or at the receiver for *L* packets.
- The communication overhead: the length of authentication information appended to each packet in order to achieve authentication.
- Delay at the sender and the receiver: delay at the sender is the number of packets that need to be processed before stream transmission and delay at the receiver is the number of packets that need to be received before authenticating the received packet.
- Resistance to packet loss: the type of loss that the scheme resists.
- Resistance to pollution attacks
- Ability to provide both authenticity and confidentiality

The comparison is undertaken for two cases: in the first case, parallelization is not used, while, in the second case, parallelization is incorporated to enhance the computation overhead. In the next subsections, results for the two cases are detailed.

A. Comparison of Sequential Multicast Authentication Protocols

Table 1 shows the comparison between Wong-Lam, PRABS, Pang *et al.* and our protocol [1]. In this table, *H* represents one hash function operation, *E* represents one erasure code function operation, *Mul* represents one modular multiplication, *Bil* represents one linear pair operation, *Add* represents one modulo add operation, *Exp* represents one modular exponentiation operation, and *Sig* represents one signature operation. Table 2 shows the computation overhead per block for the abovementioned schemes. In Table 2, the following parameters are assumed: *L* = 128 packets and assuming the use of RSA algorithm which is based on a modulus of 1024 bits. Table 3 shows the communication overhead per packet in bytes for the abovementioned schemes for different loss rate values. In Table 3, the following parameters are assumed: *L*= 128 packets, *Hout* = 16 bytes (assuming MD5 algorithm), *Sigout* = 128 bytes (assuming RSA algorithm), and *Len_P* = 128 bytes.

TABLE I. COMPARISON BETWEEN WONG-LAM, PRABS, PANG ET AL. AND OURS

	Wong-Lam	**PRABS**	**Pang et al.**	**Ours**
Computation overhead	(2m–1)H +Sig	mH+E+ (2m-1)H + Sig	m(Bil+2 Add+6Mul+Exp +2H)	mMul+ E+ mH
Communication overhead	($\log_2$m+1) * Hout + Sigout	(Hout+Sigout /m) (1+R) + ($\log_2$m+ 1)Hout	5Len_P	((Hout +4Len _P) (1+R))/ m
Delay: (assume no packet is lost) - At the sender - At the receiver	 m 0	 m m(1-R)	 0 0	 m m(1-R)
Resistance to packet loss	Any	Loss rate<R	No	Loss rate<R
Resistance to pollution attacks	Yes	Yes	Yes	Yes
Authenticity and confidentiality are provided	No	No	Yes	Yes

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE II. COMPARISON OVERHEAD PER BLOCK OF 128 PACKETS

	Wong-Lam	**PRABS**	**Pang et al.**	**Ours**
Computation overhead	255H +1024Mul	128H+ E + 255H + 1024Mul	128(Bil +Add +6Mul +Exp + 2H)	128Mul +128H +E

TABLE III. COMMUNICATION OVERHEAD PER PACKET IN BYTES FOR VARIOUS VALUES OF R

	Wong-Lam	**PRABS**	**Pang *et al.***	**Ours**
R =0.1	*256*	*147*	*640*	*5*
R =0.2	*256*	*149*	*640*	*5*
R =0.3	*256*	*151*	*640*	*6*
R =0.4	*256*	*153*	*640*	*6*
R =0.6	*256*	*156*	*640*	*7*
R =0.7	*256*	*157*	*640*	*8*
R =0.8	*256*	*159*	*640*	*8*
R =0.9	*256*	*161*	*640*	*8*

From the above tables we can conclude that:

- Our protocol consumes the lowest computation overhead among the compared protocols, as shown in Table 2. This is due to the fact that the most consuming operations to perform authentication are the signature and the multiplication operations. Assuming the use of RSA algorithm with a modulus of 1024 bits to provide authentication, the minimum number of multiplication needed to execute the RSA is 1024 multiplications. As shown in Table 2, our protocol only needs 128 multiplications to perform authentication. Therefore, our protocol has a considerable low computation overhead compared to the other protocols. On the other hand, although Wong-Lam scheme fights both packet loss and pollution attacks, it has the highest communication overhead compared to the other schemes except Pang *et al.*, as shown in Table 3.
- In addition, our protocol has the lowest communication overhead compared to other protocols (as shown in Table 3). This is due to the use of both erasure code function and a group public key. The use of group public key eliminates the need to append information for each recipient to authenticate and restore the original message, which leads to a low communication overhead. Moreover, our proposed protocol is easier to implement since both encryption and authentication are performed in one step.
- Although Wong-Lam, PRABS, Pang et al. and our protocol can fight pollution attacks, and they can authenticate any packet once it arrives, only Pang et al. and our protocol resist pollution attack and provide both confidentiality and authenticity in one step. Consequently, it could be a suitable solution for real-time applications. Moreover, our protocol can resist packet loss as opposed to Pang et al. protocol.

The next subsection presents and compares the performance of the above multicast protocols when using parallel processing.

B. Comparison of Accelerated Multicast Authentication Protocols

For fairness judgment between different protocols, they must be parallelized /(all of them must be implemented in parallel). Elkabbany and Aslan [36] introduced a parallelization technique to improve the execution time of Wong-Lam protocol. The execution time in [36] is 10% of the sequential structure of Wang-Lam protocol. Wong-Lam, Pannetrat-Molva, PRABS, and Pang et al. protocols have the same nature in terms of execution of the signature or signcryption algorithms, and then similar parallelization techniques could be applied to all of them. On the other hand, as shown in [31], Rasslan *et al.*'s signcryption algorithm which is proposed in [1, 32] achieves 70% improvement in the execution time. Although the execution time of the pipelined Rasslan *et al.*'s algorithm equals 30% of its sequential execution time, it outperforms all other parallelized protocols. Table 4 shows the computation time per packet in bytes for different protocols in case of using parallelization. In this table, assuming the use of RSA algorithm and L= 128 packets.

TABLE IV. COMPARISON OVERHEAD PER BLOCK OF 128 PACKETS AFTER PARALLELIZATION

	Wong-Lam	**PRABS**	**Pang *et al.***	**Ours**
Computation overhead	*25H +102Mul*	*38H +102Mul*	*13(Bil +Add +6Mul +Exp + 2H)*	*43Mul +43H*

V. CONCLUSIONS

Multicast protocols allow the scalable delivery of packets to a potentially unlimited number of recipients. As such, it is a very interesting mechanism for real-time applications that deliver streamed content to a large group of recipients. However, some security issues need to be solved before these applications are deployed on a large scale. The most basic needed security mechanisms for large scale commercial are confidentiality and authentication. There are different solutions to achieve authenticity and confidentiality in large scale commercial multicast protocols. Some of these solutions have drawbacks such as, high communication and computation overheads. Some other solutions are subject to packet loss and pollution attacks. Furthermore, protocols that are based on amortizing signature over several packets could be exposed to pollution attacks and inability to resist packet loss. To overcome these problems, Pannetrat-Molva and SAIDA protocols were proposed. These two protocols solve the problem of packet loss with a low communication overhead. But, these two protocols still subject to pollution attacks. On the other hand, PRABS protocol was proposed to resist pollution attacks. PRABS protocol mitigates the pollution attack problem with a high communication overhead.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Other solutions that solve the multicast authentication problem are based on signcryption techniques. Signcryption techniques aim to simultaneously accomplish the basic goals of encryption and signature schemes, that is to say confidentiality, authentication and non-repudiation. Utilization of signcryption techniques lowers the communication and computation overheads. But, due to the fact that all the block of packets must be received by the designated recipients before performing authentication, this solution cannot resist packet loss. In this paper, we suggest a protocol that is based on the idea of amortizing the signature over signcrypted texts. The proposed protocol uses signcryption technique to provide both confidentiality and authenticity and to resist pollution attacks. The proposed protocol makes use of erasure code functions to resist packet loss. The proposed protocol uses signcryption techniques to provide both confidentiality and authenticity, in one step, and lower the computation overhead. As a result, the proposed protocol has a simpler structure and easier in implementation than non-signcryption techniques. As the need for data security arises, the need to reduce the execution time and computation overhead associated with the execution of cryptographic (signcryption) algorithms increases correspondingly. In this case the need of parallelization will be a must. In this work, we utilize pipelining technique in order to reduce the computation overhead. Pipelined technique is chosen due to its suitability for signcryption algorithm nature. The pipelined technique reduces the computation time with respect to its corresponding values of a sequential execution. The proposed protocol is compared to other authentication protocols. The comparison shows that the proposed protocol has considerable low communication and computation overheads and can resist both packet loss and pollution attacks which make it suitable for real-time applications. Furthermore, the pipelined protocol is compared with other parallelized protocols. The results show that our pipelined technique outperforms the other protocols. In addition, the proposed protocol is analyzed using BAN logic to ensure that it achives the goals of authentication. The analysis shows that it achieves those goals without bugs or redundanciece.

REFERENCES

[1] H. Aslan and M. Rasslan "A New Multicast Authentication Protocol using Erasure Code Functions and Signcryption Techniques", WorldCIS-2013 Technically Co-Sponsored by IEEE, UK/IRI computer Chater, pp. 103-109, London, UK, 8 -12 Dec. 2013.

[2] K. Matsuura, Y. Zheng, and H. Imai, "Compact and flexible resolution of CBT multicast key-distribution", In Y. Masunaga, T. Katayama, and M. Tsukamoto, editors, Worldwide Computing and// Its Applications – WWCA '98, vol.1368 of Lecture Notes in Computer Science, pp. 190–205, Springer, 1998.

[3] Internet Engineering Task Force. RFC 2189: Core Based Trees (CBT version 2) MulticastRouting – Protocol Specification, 1997.

[4] L. Reyzin and R. Reyzin, "Better than BIBA: Short One-Time Signatures with Fast Signing and Verifying", 7th Australian Conference on Information Security and Privacy, pp.144-153, Melbouine, Australian, 3-5 July 2002.

[5] C. Wong and S. Lam, "Digital Signatures for Flows and Multicasts", IEEE/ACM Trans. on Networking, vol. 7, no. 4, pp. 502-513, 1999.

[6] R. Gennaro and P. Rohatgi, "How to Sign Digital Streams", 17th Annual International Cryptology Conference CRYPTO'97, pp. 180-197, Santa Brbara, California, USA, 17-21 August 1997.

[7] V. Paxon, "End-to-End Internet Packet Dynamics", IEEE/ACM Trans. on Networking, vol. 7, no. 3, pp. 277-292, June 1999.

[8] A. Perrig, R. Canetti, J. Tygar, and D. Song, "Efficient Authentication and Signing of Multicast Streams over Lossy Channels", IEEE Symposium on Security and Privacy, pp.56-73, 17 May 2000.

[9] A. Perrig, R. Canetti, J. Tygar, and D. Song, "The TESLA Broadcast Authentication Protocol", RSA CryptoBytes, vol.5, pp. 2-13, 2002.

[10] P. Golle and N. Modadugo, "Streamed Authentication in the Presence of Random Packet Loss", ISOC Network and Distributed System Security Symposium, pp. 13-22, San Digo, California, 8-9 Feb. 2001.

[11] A. Pannetrat and R. Molva, "Efficient Multicast Packet Authentication", ISOC Network and Distributed System Security Symposium, pp. 251-262, San Digo, 6-7 Feb. 2003.

[12] I. Reed and G. Solomon, "Polynomial Codes over Certain Finite Fields", J. of the Society for Industrial and Applied Mathematics, vol. 36, no.2, pp. 300-304, 1969.

[13] M. Luby, "LT Codes", 43rd Annual IEEE Symposium on Foundations of Computer Science (FOCS), pp. 271-282, Vancouver, Canada, Nov. 2002.

[14] J. Park, E. Chong, and J. Siegel, "Efficient Multicast Packet Authentication Using Signature Amortization", IEEE Symposium on Research in Security and Privacy, pp. 227-240, May 2002.

[15] C. Karlof, N. Sastry, Y. Li, A. Perrig, and J. Tygar, "Distillation Codes and Applications to DoS Resistant Multicast Authentication", ISOC Network and Distributed System Security Symposium, pp. 37-56, Feb. 2004.

[16] L. Kohnfelder, "On the Siganture Reblocking Problem in Public Key Cryptosystems", Communications of ACM, vol.31, no.19, 1995, pp.1656-1657.

[17] R. Rivest, A. Shamir, and L. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems", Communications of ACM, vol.21, no.2, 1978, pp.120-126.

[18] Y. Zheng, "Digital Signcryption or How to Achieve Cost (Signature & Encryption) << Cost (Signature) + Cost (Encryption)", CRYPTO'97, LNCS 1294, pp. 165-179, Springer-Verlag, Santa Brbara, California, USA, 17-21 August 1997.

[19] C. Li and D. Wong, "Signcryption from Randomness Recoverable Public Key Encryption", Informtion Science, vol. 180, pp.549-559, 2010.

[20] F. Li, X. Xin, and Y. Hu, "Identity-Based Broadcast Signcryption", Computer Standards and Interfaces, vol. 30, pp. 89-94, 2008.

[21] S. Duan and Z. Cao, "Efficient and Provably Secure Multi-Receiver Identity-Based Signcryption", ACISP'06, pp. 195-206, Melbourne, Australia, July 3-5, 2006.

[22] L. Pang, H. Li, L. Gao, and Y. Wang, "Completely Anonymous Multi-Recipient Signcryption Scheme with Public Verification", PLos ONE, vol. 8, no. 5, pp. 1-10, 2013.

[23] L. Chen and J. Malone-Lee, "Improved Identity-Based Signcryption", Public Key Cryptography, (PKC'2005), vol. 3386, pp. 362-379, 2005.

[24] K. Kurosawa, "Multi-Recipient Public-Key Encryption with Shortened Ciphertext", 5th Int. Workshop on Practice and Theory in the Public Key Cryptography, (PKC'2002), pp. 48-63, Paris, France, February 12-14, 2002.

[25] Z. Hu, D. Lin, W. Wu, and D. Feng, "Constructing parallel long-message signcryption scheme from trapdoor permutation", Science in China Series F: Information Sciences, vol. 50, Issue 1, pp 82-98, Feb. 2007.

[26] J. H. An, Y. Dodis, and T. Rabin, "On the security of joint signatures and encryption", In L. Knudsen, editor, Advances in Cryptology, Eurocrypt 2002, vol. 2332 of Lecture Notes in Computer Science, pp. 83–107, Springer, 2002.

[27] J. Pieprzyk and D. Pointcheval, "Parallel authentication and public-key encryption", 8th Australasian Conference on Information Security and Privacy (ACISP 2003), vol. 2727 of Lecture Notes in Computer Science, (R. Safavi-Naini and J. Seberry, editors), pp.387–401, Springer, 2003.

[28] Y.Han, X. Gui, X.Wu, and H. Yang, "Parallel Multi-recipient Signcryption for Multicast Networks", 2nd Int. Workshop on Education

Technology and Computer Science (ETCS), vol. 3, pp. 128-131,6-7 March, 2010.

[29] Y.Han, X. Gui, and X.Wu,"Parallel Multi-Recipient Signcryption for Imbalanceed Wireless Networks", Int. J. of Innovative Computing, Information and Control, ICIC , vol. 6, no. 8, August 2010.

[30] L.Zhu, F. Zhang, and S.Miao, "A Provably Secure Parallel Certificatelesss Ring Signcryption Scheme", Int. Conference Multimedia Information Networking and Security (MINES), Nanjing, Jiangsu, pp. 423-427, 4-6 Nov., 2010

[31] Ghada F. El Kabbany, Heba K. Aslan and Mohamed M. Rasslan, "An Efficient Pipelined Technique for Signcryption Algorithms", Int. J. of Computer Science Issues (IJCSI), vol. 11, Issue 1, Jan. 2014.

[32] M. Rasslan and H. Aslan, "On the Security of Two Improved Authenticated Encryption Schemes", Int. J. of Security and Networks, vol. 8, no. 4, pp. 194-199, 2013.

[33] L. Wang and C. Wu, "Efficient Key Agreement for Large and Dynamic Multicast Groups", Int. J. of Network Security (IJNS), vol. 3, no. 1, pp. 8-1, 2006.

[34] A. Yavuz, F. Alagoz, and E. Anarim, "NAMEPS: N-Tier Satellite Multicast Security Protocol Based on Signcryption Schemes", IEEE 11th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks, (CAMAD), Trento, Italy, 8-9 June 2006.

[35] M. Burrows, M. Abadi, and R. Needham,"A logic of Authentication", ACM Trans. on Computer Systems, vol. 8, Issue 1, pp. 18-36, Feb. 1990

[36] G. F. ElKabbany and H. K. Aslan,"Efficient Design for the Implementation of Wong-Lam Multicast Authentication Protocol Using Two-Levels of Parallelism", Int. J. of Computer Science Issues (IJCSI), vol. 9, Issue 3, no. 1, May 2012.

AUTHORS PROFILE

Ghada F. ElKabbany is an Assistant Professor at Electronics Research Institute, Cairo-Egypt. She received her B. Sc. degree, M. Sc. degree and Ph. D. degree in Electronics and Communications Engineering from Faculty of Engineering, Cairo University, Egypt in 1990, 1994 and 2007 respectively. Her research interests include: High Performance Computing (HPC), Computer Network **S**ecurity, Robotics, and Image Processing.

Mohamed N. Rasslan is an Assistant Professor at Electronics Research Institute, Cairo, Egypt. He received the B.Sc., M.Sc., degrees from Cairo University and Ain Shams University, Cairo, Egypt, in 1999 and 2006 respectively, and his Ph.D. from Concordia University, Canada 2010. His research interests include: Cryptology, Digital Forensics, and Networks Security.

Heba K. Aslan is a Professor at Electronics Research Institute, Cairo-Egypt. She received her B.Sc. degree, M.Sc. degree and Ph.D. degree in Electronics and Communications Engineering from the Faculty of Engineering, Cairo University, Egypt in 1990, 1994 and 1998 respectively. Aslan has supervised several masters and Ph.D. students in the field of computer networks security. Her research interests include: Key Distribution Protocols, Authentication Protocols, Logical Analysis of Protocols and Intrusion Detection Systems.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Image Zooming using Sinusoidal Transforms like Hartley, DFT, DCT, DST and Real Fourier Transform

Dr. H. B. Kekre
Senior Professor Computer Engineering Department
MPSTME, NMIMS University,
Vile Parle, Mumbai,India,

Dr. Tanuja Sarode
Associate Professor
Computer Department,
Thadomal Shahani Engg. College,
Bandra, Mumbai 50, India .

Shachi Natu
Ph.D. Research Scholar,
Computer Engineering Department
MPSTME, NMIMS University,
Vile Parle, Mumbai, India .

***Abstract*— A simple method of resizing the image using the relation between sampling frequency and zero padding in frequency and time domain or vice versa of Fourier transform is proposed. Padding zeroes in frequency domain and then taking inverse gives zooming effect to image. Transforms like Fourier transform, Real Fourier transform, Hartley transform, DCT and DST are used. Their performance is compared and Hartley is found to be giving better performance. As we increase the size of image, DCT starts giving better performance. Performance of all these transforms is also compared with another resizing technique called grid based scaling and transformed based resizing is observed to be better than grid based resizing.**

Keywords-Image zooming; DFT; Hartley Transform; Real Fourier Transform; DCT; DST

I. INTRODUCTION

Increased popularity of variety of display devices has made it necessary to adjust the image size as per observer's need with effective maintenance of image quality. Image resizing also known as image retargeting is a technique of adjusting the image into appropriate size such that its aspect ratio and salient features are preserved [1]. Resizing is required to display images on different display devices as per need. If there is a single important feature in the image then image can be cropped and resized. But when there are multiple important features in an image, resizing becomes more challenging. Displaying such images on display devices with different aspect ratios requires that all important objects must be displayed at sufficient size so that they can be easily recognized. [2]. While resizing an image, simple methods such as cropping and scaling have some obvious drawbacks. Cropping removes some image parts and is not suitable when there are multiple objects of interest as stated earlier. On the other hand, simple scaling distorts image if there is significant difference in input and output aspect ratios [3].To compromise between scaling and cropping, solution can be nonlinear data dependent scaling. Liu and Glecheir [4] and Liu H., Xie et. al [5] proposed such method for video and image retargeting. In this paper, we propose a transform based simple image resizing technique. This technique uses the relation between sampling frequency and zero padding in frequency and time domain or vice versa of Fourier transform.

II. RELATED WORK

Seam carving is the most recent method of inhomogeneously resizing the image. Seam is an optimal 8-connected monotonic path of pixels on an image from top to bottom or left to right. It is content aware image resizing technique that removes seams of low energy iteratively from the image [6]. It does not consider the global visual impact on the image while doing so. Though seam carving is better than traditional approaches like scaling and cropping, it has its own drawbacks. Since it iteratively removes or inserts low energy pixels, it may reduce visual quality of an image. It is quite possible that ROI of low energy get carved. It is time consuming method as it works pixel by pixel. K. Thilagam and S. Karthikeyan proposed piecewise seam carving method [6] in which ROIs of low energy are preserved and shape distortions are minimized. By using saliency map to automatically identify region of interest and segment the image, optimization is achieved. To adjust structure deformations, shift map editing approach is used. Another content aware image resizing method is proposed by [7]. Since human visual systems are sensitive to line structure, method proposed by authors is made line structure preserving by using similarity transforms for line structures. To control content preservation, mesh deformation is used. A shape preserving approach is proposed in [3]. In this method important local regions are passed through geometric similarity transformation and also image edge structure is preserved. To achieve this, handles are defined to both local regions and image edges and based on importance map of source image, weight is assigned for each handle.

Another popular method for resizing are using interpolation technique. Nearest neighbor interpolation [8], bilinear interpolation [9] and bicubic interpolation [10] are few of them. Nearest neighbor technique has good high frequency response but blurs the image due to aliasing. In case of bilinear and bicubic interpolation, high frequency response is not good. A grid based scaling technique has been proposed by Kekre, Sarode and Thepade [11]. It uses the area coverage of the source pixels from the applied mask in combination with their difference of their intensity for calculating new pixels values of the scaled image.

III. PROPOSED METHOD

Let f(t) be a finite time function and F(w) be the Fourier transform of given time function. If we pad zeroes at the end

Identify applicable sponsor/s here. *(sponsors)*

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

of function f(t) in time domain, sampling frequency in frequency domain is increased and thus we get more samples in frequency domain. Inversely, since forward and inverse Fourier transforms are duals of each other, if we pad zeroes at the end of transformed sequence in Fourier transform F(w), sampling frequency in time domain is also increased. When this concept is applied to digital image, it gives zooming effect to image.

Following Fig. 1 shows the end of transformed sequence for a 2-dimensional signal i.e. for a digital image when its Fourier transform is taken. Thus if we want to pad zeroes at the end of transformed sequence of image, we have to pad them such that they occupy the shaded region shown in Fig. 1. This job can be made simpler if we divide the image into four quadrants and shuffle them diagonally as shown in Fig. 2. According to Fig. 2 now we have to pad zeroes surrounding the outer boundary of image.

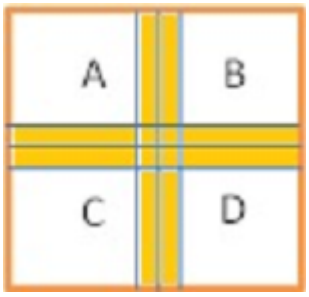


Figure 1. High frequency elements in Fourier transformed image

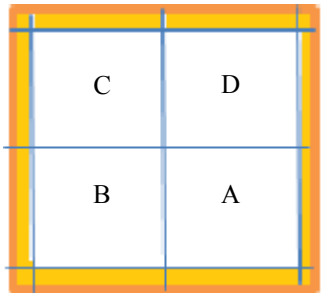


Figure 2. Diagonally shuffled quadrants of Fourier transformed image

In the proposed method different transforms like Discrete Fourier transform, Real Fourier transform [12], Hartley transform, DCT and DST are used and their performances are compared. To measure the similarity between original image quality and the resized image quality, images with increased size are reduced back to original size and then absolute pixelwise error between it and original image is found out. Smaller is the error; better is the quality of resized image. Steps of the proposed method are given below by considering the Discrete Fourier Transform.

Step 1. Read an input image.

Step 2. Take Discrete Fourier Transform of each plane to get transformed image.

Step 3. Divide transformed image into four quadrants A, B, C and D.

Step 4. Shuffle the quadrants diagonally.

Step 5. Pad zeroes to image obtained in step 4 surrounding its outer boundary.

Step 6. Reshuffle the quadrants diagonally and take inverse Fourier transform to obtained resized image.

Similar procedure can be followed to resize the image using Hartley and Real Fourier Transform because low frequency elements of these transforms are also at four corners of transformed image.

The only restriction of this method is that we can resize the image of even size to larger even size since we are dividing it into quadrants. This limitation can be overcome by repeating the central column and row in transformed image. This limitation can be overcome by using other transforms like DCT and DST. We don't need to divide an image into quadrants while padding zeroes at the end of 2-D transformed sequence i.e. DCT/DST transformed image. The reason is end of transformed sequence is located at lowermost rows and rightmost columns of transformed image as shown in Fig. 3.

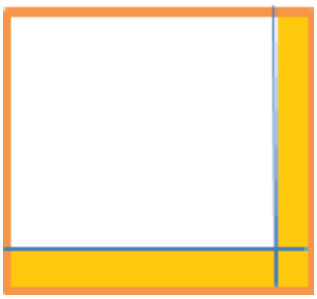


Figure 3. High frequency elements of DCT/DST transformed image

Hence padding zeroes at the lower end and at the rightmost end of transformed image will suffice. Thus division of an image into quadrants is not required and image can be resized to any size using DCT/DST.

The steps of DCT/DST based resizing are as follows:

Step 1. Read input image of size NxN and new size MxM.

Step 2. Apply DCT/DST to input image.

Step 3. Pad (M-N)xN matrix of zeroes at the bottom of image and Mx(M-N) columns of zeroes at the rightmost end of image.

Step 4. Take inverse DCT/DST to get resized image.

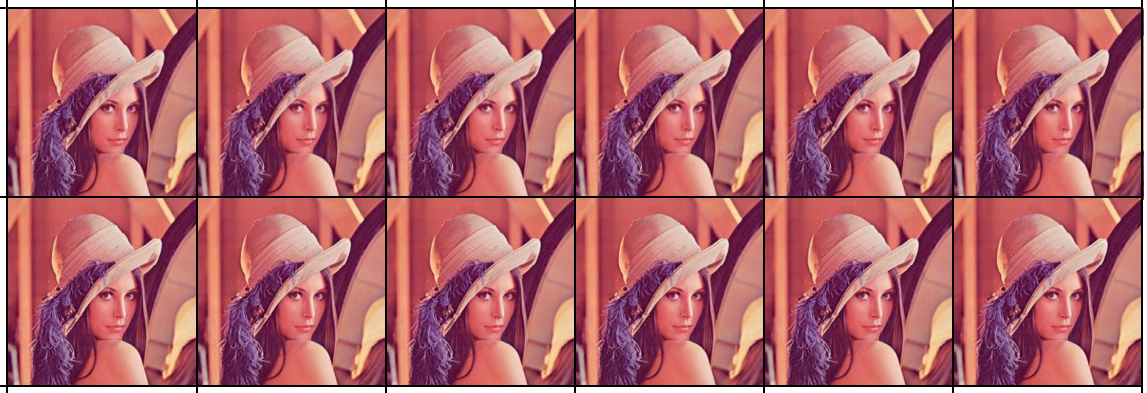
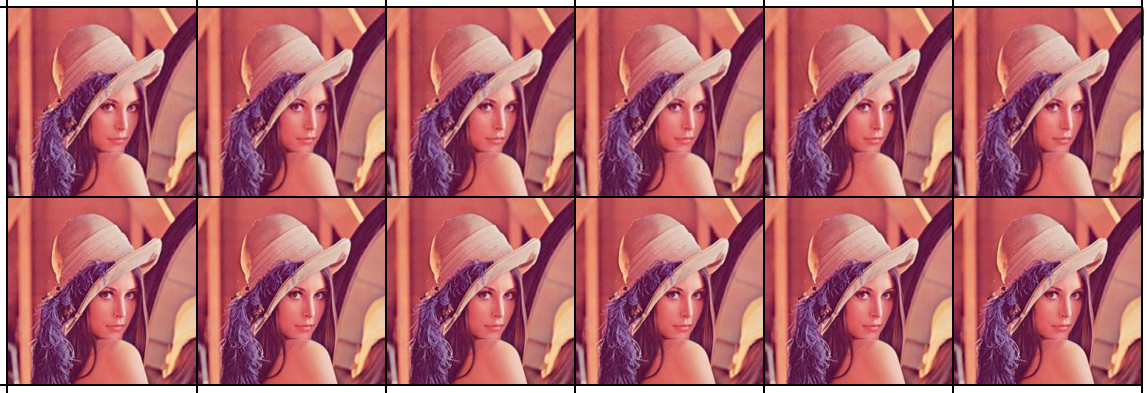
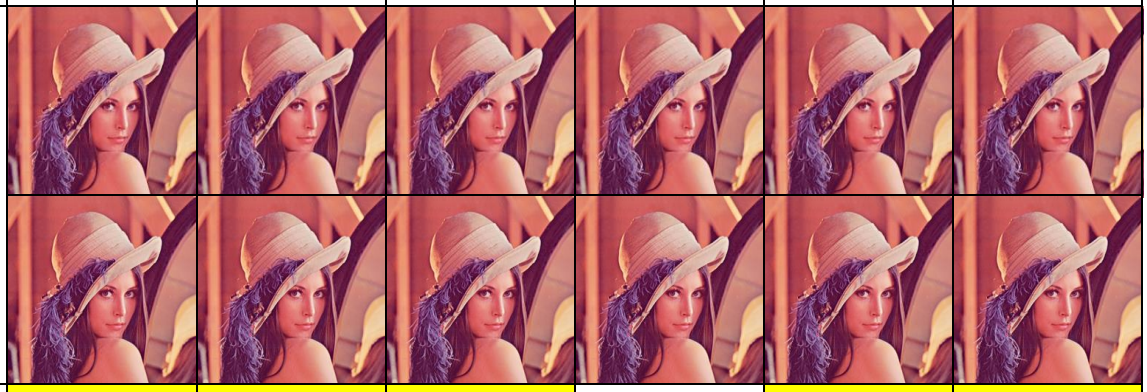
IV. RESULTS OF PROPOSED METHOD

Proposed method is applied to 'Lena' image of different sizes i.e. 128x128, 256x256 and 512x512 respectively. They are shown in Fig. 4.

Figure 4. Images used to test the proposed zooming method

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Results of proposed resizing method for Lena image of size 256x256 which is resized to different sizes (384x384, 512x512, 640x640, 768x768, 896x896, and 1024x1024) are shown in Figure 5. For each transform used for resizing, first row shows the resized images whereas image below it in the second row corresponds to the image reduced to its original size. Below every reduced image, Mean Absolute Error between original and reduced image is given to compare performance of various transforms. The last row in the table corresponds to results of Grid based resizing method.

	Resized images					
Transform used for resizing	***384x384***	***512x512***	***640x640***	***768x768***	***896x896***	***1024x1024***
DFT						
MAE between original and reduced image	0.206	0.176	0.170	0.161	0.157	0.154
Real Fourier Transform						
MAE between original and reduced image	0.218	0.177	0.155	0.142	0.133	0.127
Hartley						
MAE between original and reduced image	0.149	0.104	0.094	0.076	0.068	0.06

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

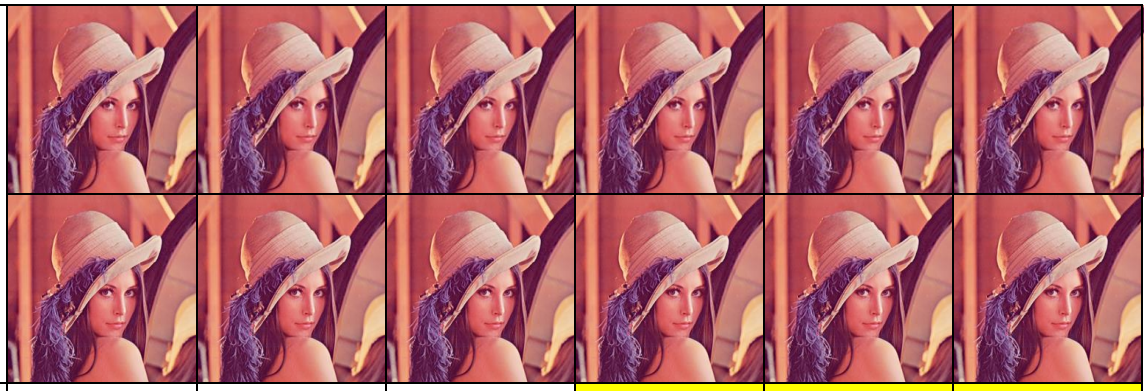
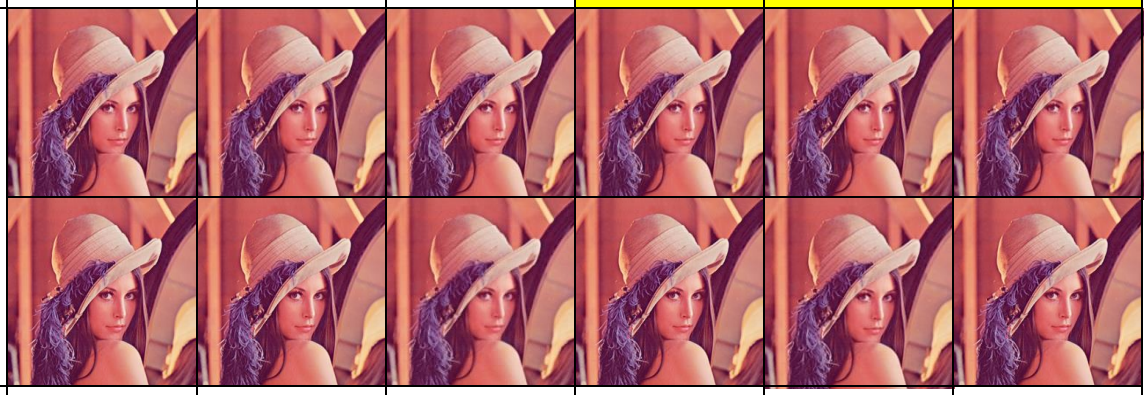
DCT						
MAE between original and reduced image	0.157	0.118	0.095	0.075	0.068	0.06
DST						
MAE between original and reduced image	0.159	0.119	0.096	0.081	0.07	0.062
GRID Based Resizing						
MAE between original and reduce image	1.702	0	0.979	0	0.676	0

Figure 5. Resized images using DFT, Real Fourier Transform, Hartley Transform, DCT and DST.

As stated earlier, comparison between the proposed resizing method using transforms like DFT, Real Fourier Transform, Hartley, DCT and DST and Grid based resizing method is done by increasing the size of an image, reducing it back to normal size and then calculating pixel wise absolute difference them.

Proposed method when applied to 'Lena' image of size 128x128 and 512x512 and compared zoomed-reduced images with original image, we get following RMSE and MAE values.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE I. PERFORMANCE OF VARIOUS TRANSFORMS USED FOR RESIZING AND GRID BASED RESIZING FOR 128X128 SIZED 'LENA' IMAGE.

Image		Lena (128x128)					
Transform used		FFT	Real FT	Hartley	DCT	DST	Grid based
Resized to							
192x192	RMSE	0.580	0.568	0.220	**0.219**	0.228	5.011
	MAE	0.446	0.258	0.154	0.163	0.164	2.857
256x256	RMSE	0.559	0.538	**0.170**	0.175	0.180	0.000
	MAE	0.433	0.225	0.111	0.123	0.124	0.000
320x320	RMSE	0.555	0.513	0.155	**0.152**	0.166	3.675
	MAE	0.429	0.203	0.100	0.100	0.103	1.663
384x384	RMSE	0.551	0.507	0.139	**0.131**	0.154	0.000
	MAE	0.426	0.193	0.083	0.080	0.088	0.000
448x448	RMSE	0.549	0.507	0.133	**0.126**	0.145	2.903
	MAE	0.424	0.186	0.075	**0.074**	0.077	1.180
512x512	RMSE	0.547	0.502	0.125	**0.119**	0.139	0.000
	MAE	0.423	0.180	0.067	**0.066**	0.070	0.000

TABLE II. PERFORMANCE OF VARIOUS TRANSFORMS USED FOR RESIZING AND GRID BASED RESIZING FOR 512X512 SIZED 'LENA' IMAGE.

Image		Lena 512					
Transform used		FFT	Real FT	Hartley	DCT	DST	Grid based
Resized to							
768x768 n back	RMSE	0.229	0.250	**0.190**	0.197	0.200	3.198
	MAE	0.179	0.163	**0.146**	0.156	0.156	1.407
1024x1024 n back	RMSE	0.187	0.208	**0.134**	0.150	0.152	0.000
	MAE	0.145	0.124	**0.102**	0.117	0.117	0.000
1280x1280 n back	RMSE	0.179	0.187	0.124	**0.122**	0.125	2.103
	MAE	0.138	0.101	**0.091**	0.094	0.094	0.811
1536x1536 back	RMSE	0.166	0.176	0.103	**0.099**	0.108	0.000
	MAE	0.127	0.086	**0.074**	**0.074**	0.078	0.000
1792x1792 n back	RMSE	0.161	0.168	0.095	**0.091**	0.095	1.558
	MAE	0.123	0.075	**0.066**	0.067	0.067	0.558
2048x2048 n back	RMSE	0.156	0.163	0.085	**0.082**	0.086	0.000
	MAE	0.118	0.067	**0.057**	0.058	0.059	0.000

From results in Table I and Table II, it can be observed that transform based resizing is better than Grid based method except when resized image is integer multiple of original image. Among the transforms used for resizing, Hartley transform gives better results closely followed by DCT, DST, Real Fourier Transform and DFT. When the size of input image is increased more and more, DCT starts functioning marginally better over Hartley. Also when input image size is more, MAE between original and zoomed-resized image is less i.e. for higher size input image proposed method's performance becomes better.

V. CONCLUSION

Proposed transform based resizing method is having less computational complexity than Grid based resizing method and hence requires less time to execute. For all transforms used for resizing, an error between original and resized-reduced image goes on decreasing as we resize the input image to higher size. Among the transforms used for resizing, Hartley transform gives better performance closely followed by DCT, DST, Real Fourier Transform and DFT in terms of MAE between original and resized-reduced image. As we resize the image to larger size, performance of DCT gets closer to that of Hartley and then performs marginally better than Hartley.

VI. ACKNOWLEDGEMENT

Authors would like to acknowledge the help rendered by Ms. Uttara Athawale and Dr. Archana Patankar.

REFERENCES

[1] Yong-Jin Liu, Xi Luo, Yu-Ming Xuan, Wen-Feng Chen, Xiao-Lan Fu, "Image retargeting quality assessment", In Computer Graphics Forum, Volume 30 Number 2, pp. 583-592, April 2011.

[2] Vidya Setlur, Saeko Takagi, Ramesh Raskar, Michael Gleicher, Bruce Gooch, "Automatic Image Retargeting", In Proceedings of the 4th international conference on Mobile and ubiquitous multimedia, pp. 59-68, December 2005.

[3] Guo-Xin Zhang, Ming-Ming Cheng, Shi-Min Hu, Ralph R. Martin, "A shape preserving approach to image resizing", Pacific graphics, volume 28, Number 7, pp. 1897-1906, 2009.

[4] LIU, F., AND GLEICHER, M. "Video Retargeting: Automating Pan and Scan", In ACM international conference on Multimedia, pp. 241–250, 2006.

[5] LIU H., XIE X., MA W., ZHANG H, "Automatic browsing of large pictures on mobile devices", Proceedings of the eleventh ACM international conference on Multimedia, pp. 148–155, 2003.

[6] K. Thilagam, S. Karthikeyan, "Optimized image resizing using piecewise seam carving", International Journal of Computer Applications, Vol. 42, No. 14, pp. 24-30, March 2012.

[7] Che-Han Chang, Yung-Yu Chuang, "A Line-Structure-Preserving Approach to Image Resizing", In Proc. of IEEE conference on Computer vision and pattern recognition, pp. 1075-1082, June 2012.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[8] H. S. Hou and H. C. Andrews, "Cubic sphines for image interpolation and digital filtering," IEEE Trans. Acoust.,Speechi Signal Processing, vol. ASSP-26, pp. 508-517, 1978.

[9] A.M. Darwish, M.S. Bedair, and S.A. Shaheen, "Adaptive resampling algorithm for image zooming," IEE Proc. Vision, Image & Signal Processing, 144, pp. 207-212, 1997.

[10] D. P. Mitchell and A. N. Netravali, "Reconstruction Filters in Computer Graphics," Computer Graphics,(Proceedings of SIGGRAPH 88), vol. 22 (4), pp. 221-228, 1988.

[11] H. B. Kekre, Tanuja Sarode, Sudeep Thepade, "Grid based image scaling technique", International Journal of Computer Science and Applications, Volume 1, No. 2, pp. 95-98, August 2008.

[12] Dr. H. B. Kekre, Dr. Tanuja Sarode and Prachi Natu, "ImageCompression Using Real Fourier Transform, Its Wavelet Transform And Hybrid Wavelet With DCT" International Journal of Advanced Computer Science and Applications(IJACSA), 4(5), 2013

AUTHORS PROFILE

Dr. H. B. Kekre has received B.E. (Hons.) in Telecomm. Engg. from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970. He has worked Over 35 years as Faculty of Electrical Engineering and then HOD Computer Science and Engg. at IIT Bombay. After serving IIT for 35 years, he retired in 1995. After retirement from IIT, for 13 years he was working as a professor and head in the department of computer engineering and Vice principal at Thadomal Shahani Engg. College, Mumbai. Now he is senior professor at MPSTME, SVKM's NMIMS University. He has guided 17 Ph.Ds., more than 100 M.E./M.Tech and several B.E. / B.Tech projects, while in IIT and TSEC. His areas of interest are Digital Signal processing, Image Processing and Computer Networking. He has more than 450 papers in National / International Journals and Conferences to his credit. He was Senior Member of IEEE. Presently He is Fellow of IETE, Life Member of ISTE and Senior Member of International Association of Computer Science and Information Technology (IACSIT). Recently fifteen students working under his guidance have received best paper awards. Currently eight research scholars working under his guidance have been awarded Ph. D. by NMIMS (Deemed to be University). At present seven research scholars are pursuing Ph.D. program under his guidance.

Dr. Tanuja K. Sarode has received M.E. (Computer Engineering) degree from Mumbai University in 2004, Ph.D. from Mukesh Patel School of Technology, Management and Engg. SVKM's NMIMS University, Vile-Parle (W), Mumbai, INDIA. She has more than 11 years of experience in teaching. Currently working as Assistant Professor in Dept. of Computer Engineering at Thadomal Shahani Engineering College, Mumbai. She is member of International Association of Engineers (IAENG) and International Association of Computer Science and Information Technology (IACSIT). Her areas of interest are Image Processing, Signal Processing and Computer Graphics. She has 137 papers in National /International Conferences/journal to her credit.

Ms. Shachi Natu has received M.E. (Computer Engineering) degree from Mumbai University in 2010. Currently pursuing Ph.D. from NMIMS University. She has 08 years of experience in teaching. Currently working as Assistant Professor in Department of Information Technology at Thadomal Shahani Engineering College, Mumbai. Her areas of interest are Image Processing, Database Management Systems and Operating Systems. She has 20 papers in International Conferences/journal to her credit.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A Self-Training with Multiple CPUs Algorithm for Load Balancing using Time estimation

Aziz Alotaibi

Department of Computer Science

221 University Ave, University of Bridgeport,

Bridgeport, CT, USA

Fahad Alswaina

Department of Computer Science

221 University Ave, University of Bridgeport,

Bridgeport, CT, USA

Abstract **- In this paper, we propose a self-trading algorithm using two new parameters: time execution and type of priority to improve the load balancing performance. Load balancing uses information such as CPU load, memory usage, and network traffic which has been extracted from previous execution to increase the resource's utilization. We have included time execution for each property individually such as CPU bound, and Memory bound to balance the work between nodes. Type of priority has been taken into account to enhance and expedite the processing of request with high priority.**

Keywords – Cloud Computing, Load Balancing, Resource allocation.

I. Introduction

The concept of cloud computing was introduced in the 1940s. And the word "cloud" has been introduced in 1990s, when the telecommunications companies' start of offer virtual private network services [1]. Cloud computing is the utilization of computer resources to deliver services over the internet. The term "cloud computing" is defined by National Institute of Standards and Technology (NIST) "Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction[2]". Users can access the cloud services through web browser or mobile application such as multi-media sharing, on-line office software, game and on-line storage. Cloud computing architecture has been divided into three services models: Software as a Service, (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). First of all, software as a service is a distribution model used by many business applications and it is available to users on the internet. Second, Platform as a service is a computing platform which allows customers to use virtualization as an integrated solution. Finally, infrastructure as a service which cloud computer has these characteristics: virtualization, reliability, scalability, performance, and security.

SaaS	End-user	Consume	Google Apps
PaaS	Developers	Build on	Windows Azure
IaaS	Operators/IT	Migrate to	Amazon EC2

Figure 1: Cloud Computing Service Levels

Recently, cloud computer technologies have developed rapidly; therefore, cloud computing problems have increased and became more complicated to solve. One of the problems that have been noticed is the high cost of using and managing super computer to cater some companies' needs. However, cloud computing has been used as solution to facilitate and utilize the computer resource and increase the performance of the system[3].

II. Load Balancing

Load balancing is the way of distributing workload across nodes to improve the resource utilization, fairness, waiting/processing delays[4]. And it is based on distributing processes on nodes to increase the performance and reduce the respond time[5]. Information collect, decision making, and data migration are the main phases that form the schema of the load balancing[6]. Load balancing policies that has

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

been used in cluster systems is either static or dynamic sachems.

III. Related Work

The main goal of using load balancer is to achieve the maximum throughput, get the response fast, and avoid overloading. In the past, There are three common algorithms proposed for load balancing [3] :

1. *Round Robin Algorithms:*

   Round Robin algorithm is based random assign a task to a node regardless of the state of the node. The disadvantage of these algorithms is that it ignores whether the system is heavily loaded

2. *Equally Spread Current Execution Algorithm:*

   Equally Spread Current Execution algorithm (ESCE) what it does is instead of giving each node a task no matter how much load it has, it assign the current task to only nodes that are free of load. The main idea is to provide equal load between all nodes.

3. *Throttled Load Balancing Algorithm:*

   Throttled Load Balancing algorithm it relays on a job manager to find appropriate the virtual machine to the job. The job manager has all virtual machine in an indexed table.

Among those algorithms, the paper has provided many analyses using Cloud Analysis Tool. The analysis was based on execution time and processing cost. The result of that analysis shows that both ESCE and Throttled Load Balancing algorithm provided (50%-60%) less execution time and less processing cost.

IV. Algorithm description

1.1 Approach

In this paper, we propose an algorithm to enhance a self-Training. Self-Tainting algorithm is tested on individual computer with a single CPU and one memory. Our idea is to replace a single CPU with multiple CPUs in an individual computer. By having more than one CPU, there will be a need to modify the algorithm suggested to keep the same efficiency level or better.

1.2 Previous Design

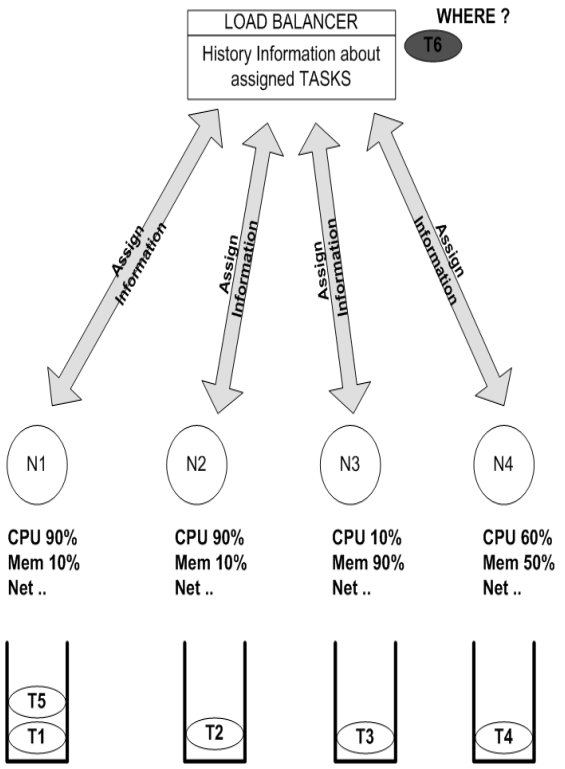


Figure 1 Previous design of the solution

1.3 Suggested Design:

When we look at the previous solution we find that when all nodes are optimized and when we have more than one candidate node for handling current received task, the previous algorithm will select arbitrary one node. The question is what if the selected candidate was taking longer time than the other not-selected candidate? So in order to optimize the algorithms, we need to redesign the solution so that it takes into account nearest node to be finished.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

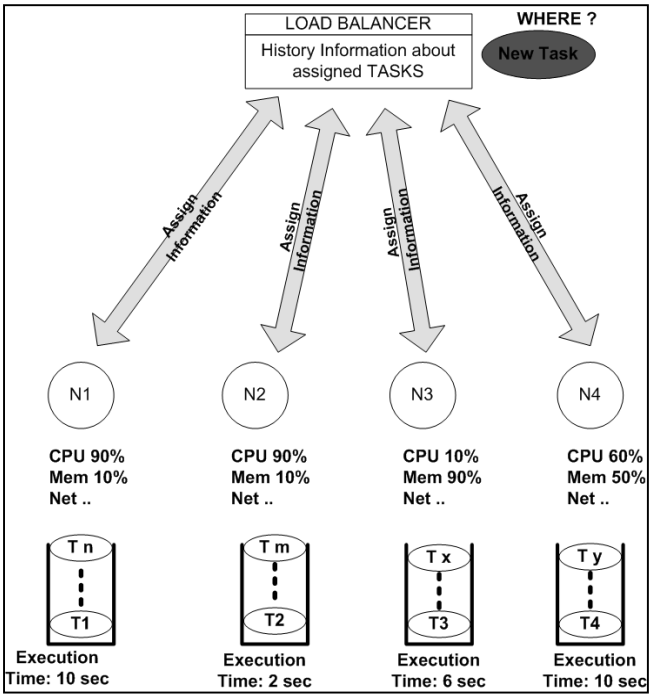


Figure 2: Not optimized solution

1.4 Design:

We have added a decision maker that we help to provide us with two things: Node information and task history.

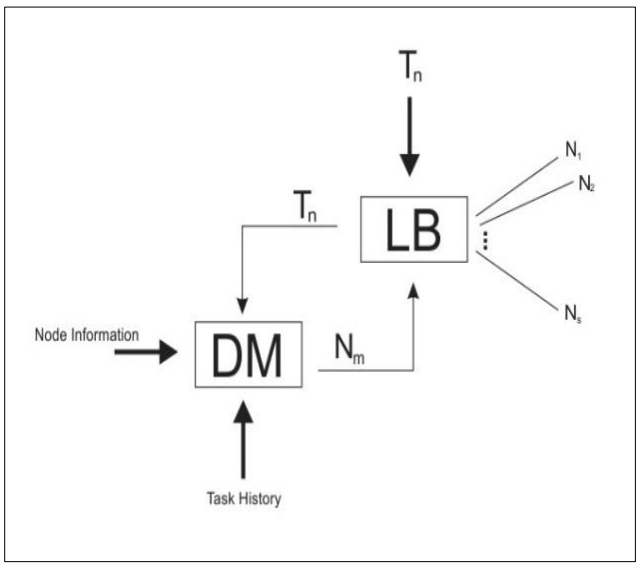


Figure 3: Communication between Load Balancer and Decision Maker

V. Task manager

Task Manager (TM) keeps the history of all tasks that has been executed and their collected information such as consumed CPU (CPUc), Memory (MEMc), Network (NETc), and Execution Time (Tc), in order to learn the task and train the algorithm. The manager stores all the information on history table. TM listens to all node managers and receives messages about current tasks being executed periodically in each node. Each time TM receives a message, CPUp for example, the manager takes the average between previous percentage and the current percentage and stores the result in the history table. Example:

$CPUc_{previous} = 20\%$ $MEMc_{previous} = 5\%$

$NETc_{previous} = 5\%$ $Tc_{previous} = 1\ \mu s$

and

$CPUc_{current} = 40\%$ $MEMc_{current} = 5\%$

$NETc_{current} = 5\%$ $Tc_{current} = 1\ \mu s$

$$CPUc = \frac{\sum CPU.previous + CPU.currnet}{N}$$

$$MEMc = \frac{\sum MEM.previous + MEM.currnet}{N}$$

$$Tc = \frac{\sum T.previous + T.currnet}{N}$$

Where N is the number of reports.

Table 1: Result of the example

Node	CPUc	MEMc	NETc	Tc
N1	30%	10%	10%	2µs
..	..	..	..	..

VI. Node manager

Each node has a node manager (NM) that monitors the node and all tasks under execution[7]. The node contains processing unit (CPU), memory unit (MEM), and network

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

link (NET). NM reports to TM all information about the task under the node such as the amount of CPU consumed, memory consumed, network bandwidth consumed, and the time spent. In order to reach the accurate measurement, NM sends the information to TM periodically (i.e. every 2 µs).

VII. Decision Making Manager

Decision Manager (DM) keeps all the information about nodes' utilization into a table (Utilization Table). The table contains node's name and overall CPU utilization, memory utilization, and the remaining time to finish.

VIII. Load Balancer:

Load Balancer (LB) performs three main tasks: receiving requests from the client, achieving best utilization, and avoiding and managing the overload[8]. Load balancer receives requests from clients and forward that to suitable node to be executed. Some information always has been attached to each process such as, CPU bound, memory bound, and network status. Node availability table and task ready queue tables are resided on the decision making device. Node availability table has information about each node and its Manager node will update the node availability tables every 30 nanoseconds continuously.

Process table:

Node	CPU bound	Memory bound	Net status
N1	40 %	90%	-
N2	73%	55 %	-

Task history table:

	CPU bound	Memory bound	Net status	Execution time
Task 1	30%	40%	-	30 ns
Task 2	70%	50%	-	40 ns
TASK 3	40%	80%	-	35 ns

Information will be combined from both tables: Node availability and history table to make a good decision. After all requests arrived to load balancer, decision making will combine all information from node availability tables, and task history table to perform best utilization.

IX. Algorithm

As shown in figure 4, whether the task is new or not, it will be directed to the making decision to assign one of the priority types.

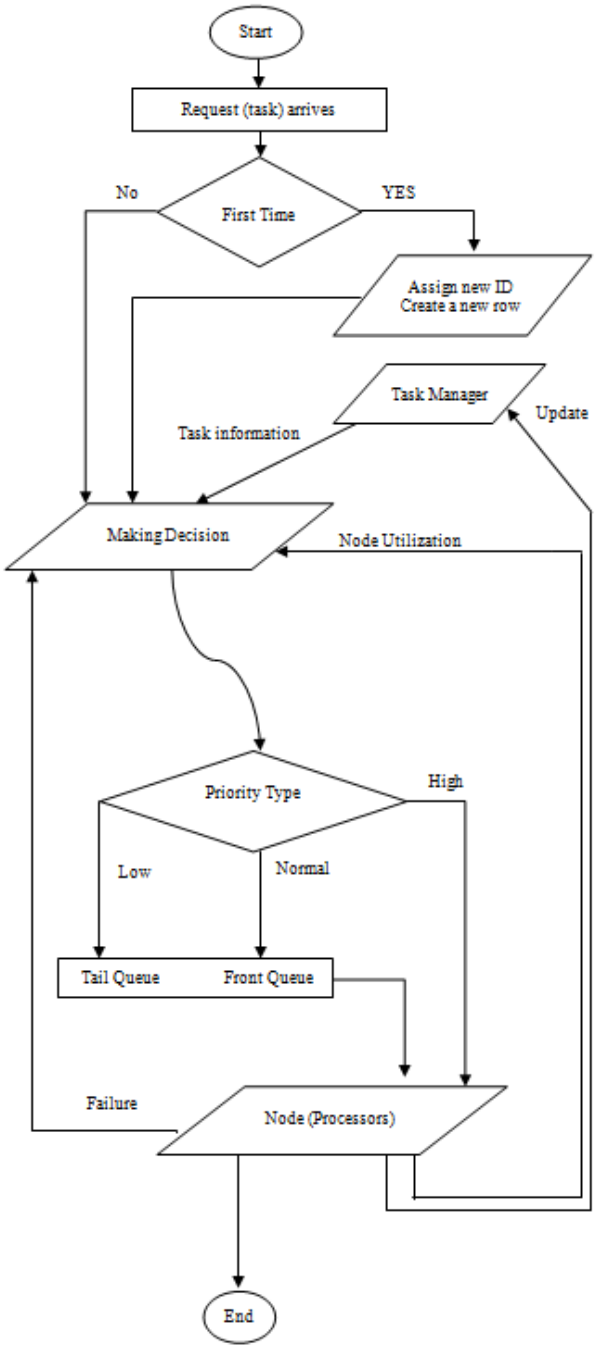


Figure 4. Our suggested algorithm

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

X. Conclusion

Cloud computing has many technologies that help in providing computation, software, data access, and storage services. User can access the data, applications, storage, and services using the browser regardless of the device and the user's location. Cloud computing has been used as solution to facilitate and utilize the computer resource and increase the performance of the system this paper illustrates our new approach that will help the load balancer to increase its knowledge about each node resources. We have included the time estimation to the previous scheme. Also, we have assigned each task one of type of priorities: low, normal, high. Priority type has been taken into account to enhance and expedite the processing of request with high priority.

References

1. Jadeja, Y. and K. Modi. *Cloud computing - concepts, architecture and challenges*. in *Computing, Electronics and Electrical Technologies (ICCEET), 2012 International Conference on*. 2012.
2. Mell, P., *The NIST Definition of Cloud Computing* September 2011
3. Randles, M., D. Lamb, and A. Taleb-Bendiab. *A comparative study into distributed load balancing algorithms for cloud computing*. in *Advanced Information Networking and Applications Workshops (WAINA), 2010 IEEE 24th International Conference on*. 2010. IEEE.
4. Andrews, J.G., et al., *An overview of load balancing in hetnets: old myths and open problems.* Wireless Communications, IEEE, 2014. **21**(2): p. 18-25.
5. Andreolini, M., et al. *Dynamic load balancing for network intrusion detection systems based on distributed architectures*. in *Network Computing and Applications, 2007. NCA 2007. Sixth IEEE International Symposium on*. 2007.
6. Tudor, D., et al. *Towards a load balancer architecture for multi-core mobile communication systems*. in *Applied Computational Intelligence and Informatics, 2009. SACI '09. 5th International Symposium on*. 2009.
7. Hung, C.-L., H.-h. Wang, and Y.-C. Hu. *Efficient Load Balancing Algorithm for Cloud Computing Network*. in *International Conference on Information Science and Technology (IST 2012), April*.
8. Mohammadpour, P., M. Sharifi, and A. Paikan. *A Self-Training Algorithm for Load Balancing in Cluster Computing*. in *Networked Computing and Advanced Information Management, 2008. NCM '08. Fourth International Conference on*. 2008.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Result-Oriented Aproach for Websites Accessibility Evaluation

Approach tested for three Dutch government websites

Marya Butt
School of Governance
Utrecht University
Utrecht, the Netherlands

Abstract— **The paper attempts to devise a result oriented approach for evaluating the accessibility of three Dutch government websites. Most of the research work pertaining website accessibility evaluation is intended to benchmark the organizations, however this study plans to initiate learning for the selected Government Bodies (GB) to improve websites accessibility. The devised approach spans three phases and is tested in three government bodies of the Netherlands. In the first phase, websites accessibility is evaluated for the selected government bodies. In the second phase, feedback from the web developers of the selected government bodies is collected to disclose their knowledge and practices. The third phase accentuates on measuring the results utilization. The websites evaluation is carried out according to the WCAG version 2.0 (level AA) by using various online tools - e.g. TAW, CCA (Color Contrast Analyzer), RIC (Readability Index Calculator) - and a test case to check that website is keyboard operable. Test results show that the selected websites failed to adhere to the WCAG 2.0. The feedback of the web developers revealed that though they are aware of these guidelines, yet clients do not want to compromise on other aspects, e.g. outlook and cost. The study initiated learning for all tested government bodies. Government bodies found the accessibility reports useful and showed perseverance to exploit research results in improving website accessibility.**

Keywords-component; E-government, Websites Accessibility, Evaluation, Netherlands, WCAG 2.0

I. INTRODUCTION

Website serves as an online tool to publicize the government information and services that are supposed to be accessible from everywhere and by everyone. Making a website that is 'accessible by everyone' is a hard nut to crack as all citizens are not equally capable i.e. citizens include people with several limitations and disabilities too and they must also be eligible to make use of website in a normal way. Extending the access to disabled people is an exigent job as disabilities constitutes hearing, visionary, speech, physical, neurological and mental disorders. Website accessibility refers 'to the extent to which maximum group of people can access the website'. Website accessibility is considered to be high if wide range of people could access the website and vice versa. Website accessibility is linked with the website design and emphasizing on website designs could enhance the website accessibility. Website design is made-up of various elements e.g. colors ratio, layout, browser independency, content (language), and support for assistive technology etc.

The International world statistics highlight the subject of websites accessibility and its importance e.g. according to World Health Organization there are above a billion disabled people in the world, and that rate is constantly rising. In the Netherlands, 22.5% of the population is reported to have some degree of physical disability. In UK there are 11.2 million disabled people. World Wide Web Consortium (W3C) is an international standard organization, comprising of member organizations, is responsible to develop World Wide Web (WWW) standards. An estimate by W3C reports that 90% of the websites, available on World Wide Web, fail to provide access to disabled people while 70-98% websites all across the world have accessibility issues. There is abundant available literature prevailing institutional web accessibility rates in various countries including UK, USA, Northern Ireland, Czech Republic, Malaysia, KSA and Oman and 90% of the studies converse about the presence of inaccessibility issues in the government websites.

The paper presents a result-oriented method to evaluate the accessibility of three government websites of the Netherlands using available online tools. The study investigates the presence of inaccessibility issues in the government websites by making use of available online tools. The latest version of Web Content Accessibility Guidelines (WCAG 2.0) is used as a benchmark for this study and online tools are exploited to provide a comprehensive picture of the website accessibility analysis. The study is conducted in three phases, at the first level tools-based websites evaluation is carried out and results-based reports are delivered to the stakeholders, at the second phase the response of the web developers of these websites is collected to reveal their knowledge and practices about the accessibility guidelines, and at the third stage measures the utilization of the evaluation carried out. During this phase, feedback (from the government body) is collected to measure the report effectiveness (evaluation utilization). The third

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

phase holds a significant value as most of the related literature about government website accessibility is intended to benchmark the government bodies by rating one government body over the other, whereas the study is an attempt to initiate learning in the government bodies to improve web design. The last phase of the research that is significantly ignored in the related literature renders the study as a result-oriented approach.

II. WEB ACCESSIBILITY GUIDELINES

Various legislations and guidelines have been developed since the issue of website accessibility has grabbed governments' attention. Table 1 provides an overview of some of the legislations regarding web content accessibility. On 3rd of December 2012, the European commission adopted a proposal for a directive on the accessibility of the public sector bodies named as 'action 64'. The objective was to make sure that by the year 2015, all the public sector websites are fully accessible. In Germany, BITV 2 came into effect on the 22nd September 2011. According to BITV 2 all the websites under the federal government, both internet and intranet, must comply with the guidelines for improving website accessibility to public. 'Stanca Act' is legislated by Italian government and it explains that the government is responsible to protect the citizens' right to avail all services and information irrespective of any disability. This law addresses public administrations including those private agencies that are licensed to work for public. Dutch accessibility law has been in effect since 2006 and according to this law, new government websites must comply with these standards however existing websites were given time to adhere to these guidelines before 2011. On 7th December 2010, BSI (British Standard Institution) launched BS 8878 first British standard to define an approach for web accessibility and it is based on the principle that the web products must be accessible to all. In USA, section 508 is a federal procurement law that stresses on all the products and services by the federal government must be accessible by everyone including people with disabilities. Section 508 is under consideration and by the year 2014 it is expected to incorporate WCAG 2.0 level AA. Apart from these countries based legislation there is an independent body W3C (World Wide Web Consortium) they developed the international guidelines (WCAG) for web content accessibility, these guidelines are very extensive and two versions of them have been released so far. The country specific guidelines mentioned in table 1 are also based on these international guidelines so WCAG serve as a superset for regional legislations. These international guidelines (WCAG) are discussed in upcoming section and are used for this study.

TABLE 1: WEB ACCESSIBILITY LEGISLATION

Law / Act	By	Based on (standard)
Action 64	European Union (EU)	WCAG 2.0
BITV 2	Germany	WCAG 2.0
Stanca Act	Italy	WCAG 1.0 (2)
Besluit Kwalitieit Rijksoverheidswesites	Netherlands	WCAG 1.0 (1)
Section 508	USA	WCAG 1.0 (1) + few additions
BS 8878 (Equality act 2010)	UK	WCAG 1.0 (2) or WCAG 2.0 (AA)

A. Web Content Accessibility Guidelines (WCAG)

WAI (Web Accessibility Initiative) by World Wide Web Consortium (W3C) in May 1999 published set of guidelines addressing website accessibility under caption WCAG (Web Content Accessibility Guidelines). The Web Content Accessibility Guidelines (WCAG) document explains how to make Web content more accessible to people with disabilities. Web "content" generally stands for the information in any Web application that may be in form of any control i.e. text, forms, images, sounds etc. So far two versions of WCAG have been released as WCAG 1.0 and WCAG 2.0. The second version is the most recent one published in 2008. WCAG 1.0 has set of 14 guidelines that have checkpoints. The checkpoints are categorized into three levels of priority namely, priority 1, priority 2 and priority 3. Priority 1 is the most important, because it covers errors which are of most important to deal with. Rest two priorities come later. WCAG 2.0 comprises of 12 guidelines which are further organized into four principles e.g. perceivable, operable, understandable and robust. 12 guidelines are divided into 61 success criteria and each criterion is assigned a specific conformance level depending upon its importance e.g. A, AA, AAA shown in table 2.

TABLE 2: WCAG 2.0 CONFORMANCE LEVELS

Standard	Conformance Level	Importance (61 Success criterion)
WCAG 2.0	A	A (priority 1) level contains 23 success criterion and they all must be fulfilled for a website to be accessible.
	AA	AA (Priority2) comprises of 13 success criterion and they should be fulfilled to improve website accessibility.
	AAA	A (Priority 3) comprise of 25 success criterion and they may be fulfilled as it would further enhance the accessibility but it is a challenge to conform to.

A website that conforms to all three levels is a big challenge therefore in all the tools and the legislated acts the highest suggested conformance level is AA. The "Conformance to a standard" means that web content satisfies the 'requirements' of that standard. The 'requirements' are the success criteria. To conform to WCAG 2.0, one need to satisfy the success criteria

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

i.e. there is no content which violates the success criteria. Moreover when it is said that website conforms to level AA it means it has passed the success criterion for level A & level AA both, same way conforming to level AAA depicts that level A, level AA & level AAA all are satisfied. Most of the tools evaluate for WCAG 2.0 level AA which is the minimum requirement for a website to be accessible to most of the group of people.

Table 3 provides an overview of what are these guidelines and how many success criterion(s) each of the guideline contains.

TABLE 3: WCAG 2.0 GUIDELINES SPAN

Web Content Accessibility Guidelines (WCAG 2.0)		
Principles	**Guidelines (12 in all)**	**Success Criterions (61)**
Perceivable	1.1 Text Alternatives: Provide text alternatives for any non-text context	1 criteria of Level A
	1.2 Time Based Media: Provide alternatives for time-based media	9 criterions: 3 of level A, 2 of level AA, 4 of level AAA
	1.3 Adaptable: Create content that can be presented in different ways e.g. simpler layout etc.	3 criterions of level A
	1.4 Distinguishable: Make it easier for user to see and hear (background and foreground)	9 criterions: 2 of Level A, 3 of Level AA, 4 of Level AAA
Operable	2.1 Keyboard accessible: Make all functionality available from keyboard	3 criterions: 2 of level A, 1 of Level AAA
	2.2 Enough time: Provide users enough time to read and use content	5 criterions: 2 of level A, 3 of level AAA
	2.3 Seizures: Do not design content to cause seizures	2 criterions: 1 of level A, 1 of level AAA
	2.4 Navigable: Make ease for user to navigate the site	10 criterions: 4 of level A, 3 of level AA, 3 of level AAA
Understandable	3.1 Readable: Make text content readable and understandable	6 criterions: 1 of level A, 1 of level AA, 4 of level AAA
	3.2 predictable: Make web pages appear in predictable ways	5 criterions: 2 of level A, 2of level AA and 1 of level AAA
	3.3 input assistance: Help user avoid mistakes	6 criterions: 2 of level A, 2 of level AA and 2 of level AAA
Robust	4.1 Compatible: Maximize compatibility with current and future user agents	2 criterions of level A

B. Website Accessibility Evaluation Tools

Various online tools are available to evaluate web content accessibility and some sites have listed them according to their specifications and benchmark they meet. Various tools are free to use and few of them have options to evaluate across multiple standards e.g. WCAG, section 508, and Stanca act. WCAG option is available in all tools because they are internationally recognized guidelines and all the other legislations (guidelines) are based on them. Two versions of WCAG have been released so far and as the second version is the latest one so most of the tools evaluate according to WCAG version 1.0, and few tools have option to evaluate accessibility adherence for WCAG 2.0.

TABLE 4: WEBSITE ACCESSIBILITY TOOLS

Tool	**Available Testing options (Standards)**
Total Validator	WCAG 1.0 Section 508
A-CHECKER	WCAG 1.0 **WCAG 2.0** Section 508 Stance Act BITV 2
WAVE	WCAG 1.0 Section 508
TAW	WCAG 1.0 **WCAG 2.0**
Accessibility Check	WCAG 1.0
WAEX	WCAG 1.0

Table 4 provides an overview of few online tools that could be exploited for websites accessibility evaluation, and it is shown that only two tools i.e. TAW and A-Checker have option for WCAG 2.0. TAW tool is available in English and Spanish interfaces and it generates the evaluation results in form of errors, warnings and 'not reviewed'. Not reviewed are the accessibility issues which TAW failed to analyze and they requiring human review. A-Checker generates a report with highlighting the error and its possible repair suggestion. As all tools have their own algorithms in which they are built so it is probable that one tool detect some errors while other (tool) detects some others. However all tools are likewise and easy to operate. For evaluation of any website, its URL is provided and the tool generates the website accessibility results.

Apart from the tools mentioned in table 4, there are some other tools that evaluate across a single principle or a guideline e.g. Readability Index Calculator (RIC) evaluates the difficulty level of the website content (text) according to WCAG. As a government website is used and visited by all sorts of people and therefore the language used in website is of vital importance. This tool is available in different languages e.g. French, German and Dutch etc. It operates by copying the text from the website into the RIC interface and it generates the difficulty level of that text.

Color contrast analyzer (CCA) calculates the contrast ratio between the foreground and the background colors used in the website on the basis of Web Content Accessibility Guidelines as recommended by W3C. For a text to be visible and readable it must be in adequate contrast to its background e.g. white text on white background is not visible however black text on white background or vice versa is readable because of the high contrast ratio. WCAG has defined the minimum ratio required for both small and large texts i.e. for large text >=3:1 and for small text >=4.5:1. To calculate the contrast ratio, besides CCA there are many other tools available that could be

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

exploited e.g. AccessColor, Accessibility Color Wheel, and Color Laboratory etc.

Synopsis: This section showed that there are multiple tools available to evaluate the website accessibility across various standards and any single of them or some of them in combination could be exploited to highlight the maximum areas where improvement could enhance the website accessibility.

III. RESEARCH METHODS

The paper deals with evaluating the website accessibility for three governing bodies of the Netherlands. The selection of the websites is made on the fact that these government bodies got their clients from all age groups and belong to different streams of public sector i.e. one is a large sized, second is medium sized and the third one is public university. The whole process of websites accessibility evaluation is expected to be completed in three steps as follows.

1. Tool based evaluation of selected websites, report preparation and delivered to the concerned department for further action.
2. Response from the web developers of the selected websites to reveal their knowledge about the accessibility guidelines and practices.
3. Feedback about the delivered report, how useful it was? Did the evaluation findings initiate learning process?

At the first stage, all three government bodies are analyzed across five tests i.e. TAW, AChecker, CCA, RIC, and a test case. TAW and AChecker both evaluate across WCAG 2.0 and reason being using both for the research is that both have their own algorithms and therefore their combination could increase the possibility of revealing and highlighting more areas where improvement could increase the website accessibility. The results of all the tools are compiled into a report which is delivered to the concerned department of that public body for consideration.

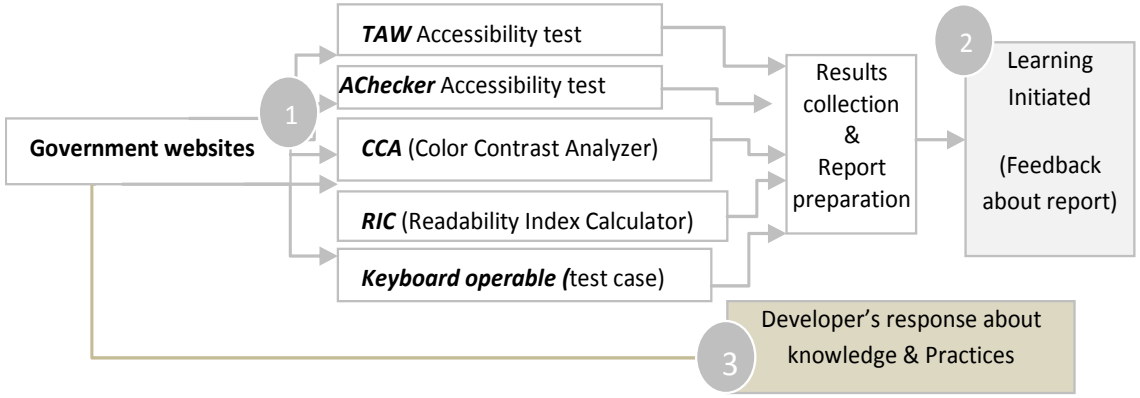


FIGURE 1: EVALUATION PROCESS FLOW

At the second phase, web developers of the selected websites are asked concerning their familiarity and knowledge about the websites accessibility guidelines shown in fig. 2. The phase seeks to reveal the reasons behind existing inaccessibility issues in the websites.

- Main feature that client emphasizes on, while designing website (e.g. content, navigability, readiness, accessibility, color's scheme)".
- Developers familiarity with the accessibility guidelines mentioned in Dutch law
- Developers familiarity with the WCAG
- Importance that clients give to follow standards
- Developer's special education and training regarding WAG (Web Accessibility Guidelines)
- Developers test website across any tool or any assistive technology
- Priority that developers give to these accessibility guidelines while designing websites
- Awareness regarding availability of different online tools for evaluating websites

FIGURE 2: WEB DEVELOPER'S PRACTICES & KNOWLEDGE

At the third phase, feedback from the government bodies (web department) is collected to measure the usefulness of the accessibility report. Feedback form is based on four parameters identified in program evaluation standards i.e. utility, feasibility, propriety, and accuracy shown in fig. 3. A simple questionnaire is designed with closed ended questions based on expanding these four parameters with answering options i.e. 'met, 'partially met', and 'not met'.

Utility	**Propriety**
- The report findings are impactful - Findings are informative and beneficial - Addressing the findings could cause improvement - Findings are in interest of stakeholders - The rationale and procedures mentioned in the report are easy to interpret - Findings are credible enough to be accepted - Report disseminated to right people	- Findings don't violate any legal matters - Findings are ethically sound - Findings are easy to comprehend - Findings does not address to conflict of interest of any one **Feasibility** - Findings are viable to implement **Accuracy** - Information in the findings is technically right - Findings are valid to be conceded Findings are not biased

FIGURE 3: WEBSITE ACCESSIBILITY REPORT EVALUATION

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

IV. RESULTS

The accessibility testing of all three websites are evaluated across TAW and AChecker. The testing generates the output in three categories i.e. errors, warnings and "couldn't reviewed". Errors are the "problematic areas (seriously) required to be fixed", warnings implies for "potential errors" they may lead to errors so it's better that they should also be addressed as well , however there are few criterions which the tools could not review and they require human review for testing, come under caption " could not reviewed". The summary of the results is shown in table 5. Table 5 depicts that according to WCAG 2.0, GB 1 website contains 4 errors (problems), 127 warnings and 17 'could not reviewed' issues. GB 2 website has 11 problems, 83 warnings, and 16 'could not reviewed' issues. GB 3 website holds 28 problems, 193 warnings, and 17 'could not reviewed' issues.

TABLE 5: RESULTS GENERATED BY TESTING

Governing body(GB)	Tools used	Results		
		Problems X	Warnings !	Could not Reviewed ?
GB 1	**TAW**	4	127	17
GB 2	**ACHECKER**	11	83	16
GB3		28	193	17

Table 6 provides an expanded view of the errors and warnings, illustrating, in which WCAG principle each error or warning actually exists for all government bodies (GB). According to table 6, under 'perceivable' principle GB 1 contains 4 problems and 31 warnings, GB 2 has 7 errors and 23 warnings, and GB 3 engrosses 24 errors and 72 warnings. For principle 'operable' GB1 shows 28 warnings, GB 2 shows 4 errors and 37 warnings and GB 3 contains 2 errors and 44 warnings. For third principle 'understandable' GB 1 holds 6 warnings, GB 2 has 0 error and 6 warnings, and GB 3 contains 1 error and 6 warnings. Under last principle 'robust' GB 1 has 62 warnings, GB 2 contains 17 warnings, and GB 3 contains 1 error and 71 warnings.

TABLE 6: EXPANSION OF ERRORS & WARNINGS

GB	Testing analysis (WCAG 2.0)			
	Perceivable	Operable	Understandable	Robust
GB 1	4 **(X)** 31 **(!)**	0 **(X)** 28 **(!)**	0 **(X)** 6 **(!)**	0 **(X)** 62 **(!)**
GB 2	7 **(X)** 23 **(!)**	4 **(X)** 37 **(!)**	0 **(X)** 6 **(!)**	0 **(X)** 17 **(!)**
GB3	24 **(X)** 72 **(!)**	2 **(X)** 44 **(!)**	1 **(X)** 6 **(!)**	1 **(X)** 71 **(!)**

There are some success criterions in WCAG 2.0 that require human review e.g. guidelines relating to color contrast, readability, support for assistive technology. It is better to deal them separately in accordance with the guidelines. All three websites are tested for the use of sufficient color contrast ratio in the website in line with the standard set by WCAG 2.0. Table 7 shows the results generated by CCA (Color Contrast Analyzer) for all three government websites. GB 1 and GB 2 failed to hold minimum contrast ratio standard by WCAG. GB 3 cleared the test with high contrast ratio of 10:1, the high the contrast ratio the more visible is the text.

TABLE 7: RESULTS OF CCA

GB	Tool	Standard ratio (WCAG)	Ratio used in the website	Remarks (Fail/Pass)
GB 1	**Color contrast analyzer (CCA)**	For small text> 4.5:1 For large text >3:1	1.9:1	X
GB 2			2.3:1	X
GB 3			10:1	√

To check the difficulty level of the content used in the selected government websites, RIC (Readability Index Calculator) is used. Text from all three websites is pasted into the interface of the RIC and it operates by revealing the difficulty level of the language used in these websites. Besides English language, RIC tool supports Dutch language text. Table 8 shows the results with revealing the difficulty level of the language used in the selected websites e.g. according to the legend of RIC GB 1 contains fairly difficult content, GB 2 holds content that is normal and the content of GB 3 according to RIC is difficult to comprehend.

TABLE 8: RESULTS OF RIC

GB	Tool	Index calculated	Remarks	Legend
GB 1	**Readability index calculator** (RIC)	50	Fairly difficult	90-100 Very Easy 80-90 Easy 70-80 Fairly Easy 60-70 Normal 50-60 Fairly Difficult 30-50 Difficult 0-30 Very Difficult
GB 2		65	Normal	
GB 3		31	Difficult	

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

There are two success criterions regarding keyboard accessibility in "operable" principle of WCAG 2.0. According to WCAG the website must be fully operable using keyboard only and there must be no keyboard trap. To check the website accessibility through keyboard, a simple test case is conducted. In the test case, all three websites are traversed with only keyboard and 200 operations are performed, the results are notified.

TABLE 9: TEST CASE FOR KEYBOARD OPERABILITY

GB	Number of keyboard hits	Accomplished (success rate)	Remarks
GB 1	200	99%	Keyboard operable, no trap
GB 2	300	100%	Keyboard operable, no trap
GB 3	200	100%	Keyboard operable, no trap

V. DISCUSSION AND ANALYSIS

This section introduces the existing issues in each website along with their relevance to accessibility. Table 10 shows the error types of GB1, two types of errors exist and each error has two instances making 4 errors in total.

TABLE 10: TYPES OF ERRORS FOR GB 1

Conformance level(PL)	Error type	Instances	Intent
A	Consecutive text and images link to same resource	2	The objective is to avoid unnecessary duplication due to presence of adjacent text and iconic versions of a link.
AA	Use of absolute font sizes	2	The objective of this technique is to identify and specify the font size of text proportionally so that user could efficiently scale the content.

According to WCAG 2.0, if consecutive links leads to same resource file, they should be grouped; this rule is violated in GB1 design as the same links are not grouped. And the second error is regarding the usage of absolute units for font sizes which should not be done because else wise it does not allow users to scale (big or small) the content effectively.

On the other hand in GB 2 website accessibility test highlighted 11 errors in all. There are four types of errors but their multiple instances make 11 errors in total. Table 11 provides a list of these four errors and their occurrences (instances) in GB 2 website design.

TABLE 11: TYPES OF ERRORS OCCUR IN GB 2

GB 2 Level	Error type	Instances	Intent
A	Images without "alt" attribute	1	"The intent of this success criterion is to make information conveyed by non-text content accessible through the use of a text alternative".
A	Form controls without label	4	"If non-text content is a control or accepts user input, it should have a name or label that describes its purpose".
AA	Use of absolute font sizes	2	"The objective of this technique is to identify and specify the font size of text proportionally so that user could efficiently scale the content".
A	Empty links (navigation)	4	"Whenever possible, provide link text that identifies the purpose of the link without needing additional context".

According to WCAG all the images, that convey meaning and provide understanding of the content, must be provided with a short 'alt' attribute. This enhances the website interactivity as alt attribute is displayed when the element cannot be rendered normally. The second error is the 'presence of form controls without associated labels', according to WCAG use 'label' elements or 'title' to label the form control. When there is no text on the screen that could be identified as label or in case where it is confusing to display label, user agent (software that works on user behalf) can speak the title attribute. The third error is same as identified in GB 1 for using absolute font sizes and the forth error is the presence of empty links in the website. According to WCAG a link should never be empty, it should always contain text else it could cause confusion for keyboard or screen reader users.

Table 12 provides an overview of the errors occur in GB 3 website. Six types of errors with multiple instances are identified, 3 errors are same as highlighted in GB 1 and GB 2. According to WCAG the headings in the webpage must be defined in different level headers e.g. h1, h2 etc. Using headings merely confuse users, who rely on them for navigation or for those using assistive technologies and in GB 3 website none heading level is defined. The second error emphasizes to follow a standard method for the form submission, according to WCAG a form must have a submit button as it is an appropriate control for causing change of context. Third error indicates the absence of well-formedness of the webpage it implies that web content must be robust enough to be interpreted by user agents and assistive technologies.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE 12: TYPES OF ERRORS OCCUR IN GB 3

Priority level(PL)	Error type	Instances	Intent [19]
A	Consecutive text and images link to same resource	2	The objective is to avoid unnecessary duplication due to presence of adjacent text and iconic versions of a link.
A	None h1 element in the document	2	The intent is to organize the content
AA	Use of absolute font sizes	20	The objective of this technique is to identify and specify the font size of text proportionally so that user could efficiently scale the content.
A	Empty links (navigation)	2	"Whenever possible, provide link text that identifies the purpose of the link without needing additional context".
A	Form with no standard submission method	1	The objective of this guideline is to allow user to explicitly request changes of context.
A	Web page is not well formed	1	Content must be capable of interpreted by variety of user agents including assistive technologies.

Apart from the accessibility testing tools, three independent tools e.g. CCA, RIC and a test case was conducted mentioned in section 5. Table 13 provides the summary of the results for all three government bodies.

TABLE 13: WEBSITE ACCESSIBILITY TESTS SUMMARY

Benchmark	Testing issues	Tools	Results GB 1	GB 2	GB 3
WCAG	Perceivable	TAW AChecker	X	X	X
	operable		√	X	X
	understandable		√	√	X
	robust		√	√	X
	Color contrast	CCA	X	X	√
	Readability	RIC	X	√	X
	Keyboard trap	Test case	√	√	√

Table 13 shows that only GB 3 cleared color contrast test while other both websites failed the test as the contrast ratio for the colors used in background and foreground (of the websites) is not high enough to make it readable for everyone e.g. light background color with light foreground color. GB3 passed this test as the contrast ratio between the used colors is very high making text readable e.g. light background with dark foreground text color or dark background with light foreground text color. Readability test was only passed by GB2 that contains text which is normal to understand while other two websites content is difficult to comprehend as identified by the tool.

After performing series of tests for the selected government bodies, it is observed that all three government bodies don't fully adhere to the international standards of web content accessibility guidelines set by W3C. GB 1 and GB 2 both cleared 4 tests out of total 7 while GB 3 could only clear 2 tests.

The feedback from the web developers of these websites was collected via email across the issues mentioned in section 4. The gist behind obtaining their response is to reveal their practices and knowledge about the accessibility guidelines. The form was mailed to all three web development companies and response was received from GB 1 and GB 2 while GB3 refused to reply.

The response of the website developers throws light on the fact that clients usually emphasize on outlook, content, navigability and colors scheme parameters. Both web developers (according to their response) are proverbial with the international web content accessibility standards as well as with the Dutch web content accessibility law, despite of the fact that none of them have got any special training in that sector. GB1 developer tries to conform to international web standards and is familiar with the available testing tools while GB2 developer tries to follow Dutch standards and don't conduct any test to check the conformity level. According to GB1 developer, their importance for following these guidelines is connected to the user demand, as they can't enforce the clients to make them follow the standards. On inquiring "whether they run any testing while designing websites", GB 1 developer evaluates using online tools for accessibility concerns, while GB2 developer undergoes usability testing (along with normal users) which is not sufficient as user solicited test does not identify any accessibility issue which is our point of concern, that supports the reason why GB1 overall showed fewer errors than GB2. GB1 developer considers all the guidelines difficult to conform and according to him the test generates lots of errors and therefore he feels "need of improvement" in that area, which is practically not possible as all the guidelines (level AA) are designed by W3C keeping every possible user group in account and neglecting some or more guidelines means depriving some people from accessibility. GB1 web developer response also establishes the fact that the developer's familiarity with these guidelines is not adequate enough to grasp the gist behind pursuing those guidelines in website designs.

VI. MEASURING RESULTS UTILIZATION

The available literature regarding websites accessibility evaluation is mostly intended to benchmark the government bodies however this study is result oriented with an attempt to initiate learning for the government bodies. The real worth of an evaluation resides in its results utilization, however this phase is neglected in most of the related literature. To measure results utilization, reports from the test results along with the suggestions are prepared and delivered to the web department

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

of the government bodies. The task doesn't end after just delivering the report, response about the report effectiveness is gathered after a month it is delivered.
For response collection about the report, a form is constructed by expanding the parameters identified in program evaluation standards mentioned in section 4.

TABLE 14: ACCESSIBILITY REPORT EFFECTIVENESS

Government Bodies (GB)	Report evaluation parameters			
	Utility	Feasibility	Propriety	Accuracy
GB 1	7/7 100% met	1/1 100% met	3/4 met, 1/4partially met 75%	3/3 100% met
GB 2	100% met	100% met	100%met	100% met
GB 3	100% met	100% met	100%met	100% met

The forms were delivered to the web administration department of the government bodies to measure the effectiveness of the report. The response obtained about the reports utility was positive and constructive with none of the parameter marked as 'did not met'. Table 14 shows the usefulness of the report as interpreted by the web department of the government bodies i.e. indicator 'utility' is expanded into seven questions and all three government bodies marked the seven criterions of utility to be 'met'. 'Feasibility' contains one question and according to the all government bodies, report 'met' the feasibility criteria. Third indicator 'propriety' is synthesized into four questions, GB 2 and GB 3 found all four criterions to be 'met' while GB 1 consider one criteria to be 'partially met' and rest three as 'met'. The last indicator 'accuracy' spans three questions and all government bodies marked all three as 'met'. The government bodies showed perseverance to exploit research results in enhancing website accessibility.

VII. CONCLUSION

The result oriented approach that was followed by measuring the usefulness of the evaluation results, initiated the learning in the government body as they welcomed the testing results. Most of the related literature is missing this part and is restrained to only testing results and benchmarking the government bodies. However as identified by various researchers, the real worth of any evaluation based studies is gauged by its effective utilization.

The study reveals that all three websites still require considerable efforts in their designs to make them fully accessible. Web developers seems to give priority to the guidelines but their services are clients driven who usually don't want to make any compromise on the other aspects like visual outlook and budget to give extra time to follow guidelines.

It is observed that the web developers and administrators could easily exploit the tools, which are available online and are easy to operate. The effective use of these tools could significantly improve the websites accessibility along with saving money and time both.

The reports generated by the tools helps to identify the errors and facilitate to fix them. Every accessibility evaluation tool operates on its algorithm which differs from the other tool and hence each tool has its own limitations depending on the algorithm used. It is therefore better to use combination of tools to identify maximum accessibility issues.

Response of the web developers showed that only will of the developer does not matter, there must be same attitude, stipulation and emphasis on adherence to standards from developers and government bodies both so that website could be design on the guidelines legislated by the Dutch government hence ensuring the e-government services delivery on equitable basis. Laws are there but the need is to enforce their implementation.

REFERENCES

[1] M. Paris, "Website accessibility: a survey of local e-government websites and legislation in Northern Ireland", Universal access in the information society, 4(4), 2006, pp. 292-299

[2] X. Zeng, "Evaluation and Enhancement of Web Content Accessibility for Persons with Disabilities", *Doctoral Dissertation*, Graduate Faculty of School of Health and Rehabilitation on Science in partial fulfillment of the requirements for the degree of Doctor of Philosophy, University of Pittsburg, 2004,
http://etd.library.pitt.edu/ETD/available/etd-04192004-155229/unrestricted/XiaomingZeng_April2004 .pdf

[3] World report on disability, World Health Organization, 2011,
http://whqlibdoc.who.int/publications/2011/9789240685215_eng.pdf

[4] ANED, Academic Network of Experts on Disability, The Netherlands – ANED country profile, Human European Consultancy, University of Leeds, 2009
http://www.disability-europe.net/content/pdf/Netherlands%20ANED%20country%20profile.pdf

[5] FRS, "Disability prevalence estimates", It is based on data taken from the Family Resources Survey, 2011
http://odi.dwp.gov.uk/docs/res/factsheets/disability-prevalence.pdf

[6] C. Boldyreff, "Determination and Evaluation of Web Accessibility", Paper presented at the Eleventh IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE'02), Pittsburgh, PA, USA, 2002

[7] J. Lazar, A.D. Sponaugle, and K.D. Greenidge, "Improving web accessibility: a study of webmaster

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

perceptions", Computers in Human Behavior, 20(2), 2004, pp.269-288.
[8] J. M. Kuzma, Accessibility design issues with UK e-government sites, Government information quarterly, 27(2), 2010, pp.141-146
[9] S. McCoy, and E.T. Loiacono, "Website accessibility: a cross-sector comparison", Universal access in the information society, **4**(4), 2006, pp.393-399
[10] H. Kopackova, K. Michalek, and K. Cejna, "Accessibility and findability of local e-government websites in the Czech Republic", Universal access in the information society, **9**(1), 2010, 51-61
[11] M. H. A. Latif, and M. N. Masrek, "Accessibility Evaluation on Malaysian E-Government Websites", Journal of E-Government studies and best practices, 2010, pp. 1-11
[12] A. Abanumy, A. Al-Badi, and P. Mayhew, "e-Government Website Accessibility: In-Depth Evaluation of Saudi Arabia and Oman", The Electronic Journal of e-Government, 3 (3), 2005, pp. 99-106.
[13] Europa, Digital Agenda: Commission proposes rules to make government websites accessible for all. 2012 [Press Release].
http://europa.eu/rapid/press-release_IP-12-1305_en.htm
[14] M. Batusic, "Guidelines and directives on accessibility", Fabasoft, eGov-Suite, Rethinking e-government. 2012, April 13
http://blog.egov-suite.com/2012/04/guidelines-and-directives-on-accessibility/
[15] R. Caffo, "Exchange of good practices and policies: accessibility standards for people with disabilities: Italian good practices", 4° Meeting of the Member States' Expert Group on digitisation and digital preservation Luxembourg. 2009,
http://ec.europa.eu/information_society/activities/digital_libraries/doc/mseg/meetings/4th/our_presentations/it_accessibility.pdf
[16] S. K. Murphy, "WCAG 2.0: Emerging International Standard for Web Accessibility and Video Captions", 2013,
http://www.3playmedia.com/2013/11/25/wcag-2-0-emerging-international-standard-web-accessibility-video-captions/
[17] Nomensa, "BS8878 Web Accessibility Code of Practice", Nomensa Blog. 2010, December 10
http://www.nomensa.com/blog/2010/bs8878-web-accessibility-code-of-practice/
[18] W3C, Web Content Accessibility Guidelines *2*.0, (2008), http://www.w3.org/TR/WCAG/
[19] M. Q. Patton, "Utilization-focus evaluation", 3rd edition, Thousand Oaks, CA: Sage. 1997.
[20] Joint Committee on Standards for Educational Evaluations, The Program Evaluation Standards: How to Assess Evaluations of Educational Programs, 1994. Thousand Oaks, CA: Sage Publications.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Performance Evaluation of Forward Difference Scheme on Huffman Algorithm to Compress and Decompress Data

Adamu Garba Mubi
Computer Science Department
Federal Polytechnic, Mubi.
Adamawa State, Nigeria

Dr. P. B. Zirra
Computer Science Department
Federal University Kashere
Gombe State, Nigeria

Abstract - Data Compression using Forward Difference Techniques on Huffman algorithm is a research work which investigated how Forward Difference Techniques was used on Huffman to compress and decompress data without loss of information. The study measured the performance of Huffman algorithm against the Forward Difference on Huffman using Compression Ratio, Compression Factor and Saving Percentage. During the encoding the new algorithm reads the input file, serializes the distinct characters, determines the probability of each character, computes Forward Difference on the positions of each character, computes twos complement on the resulting difference, computes the new probability using the twos complement code, determines the codeword for each distinct character and finally determine the binary symbols to be transmitted. While decoding the new algorithm reads the whole encoded message bit-by-bit, determines a codeword from the coded message and determines a symbol the codeword represented; using the new probability the twos complement code is regenerated. Decimal equivalent of the twos complement described a delta difference. Backward difference is used to determine the character positions of each character which is used again to reconstruct the whole message file. The results obtained revealed clearly that the performance of Forward Difference on Huffman is better than that of Huffman alone.

Keyword: Data, Huffman algorithm, data compression

I INTRODUCTION

Computers process miscellaneous data [7]. Some data, such as pictures, voices or videos, are analogue. Present-day computers do not work with infinite-precise analogue values, so the data has to be converted to a digital form. Digitization, according to

[1] added that it is a process by which text, images, audio and video are converted to the digital form. Once the data is in the digital form, it can be converted to another digital form without loss of quality; unlike the analog types which degrades with each use and losses quality when copied, [13]. That is, the infinite number of values is reduced to a finite number of quantized values. Therefore some information is always lost (i.e. lossy type).

Regardless of the way data are gathered to computers, they usually are sequences of elements. The elements come from a finite ordered set, called an alphabet. The elements of the alphabet, representing all possible values, are called symbols or characters. One of the properties of a given alphabet is its number of symbols, which is called the size of the alphabet. For a Boolean sequence the alphabet consists of only two symbols: false and true, represented as 0 or 1 bit only, [7].

[10] further added that a sequence of symbols can be stored in a file or transmitted over a network in a compressed form, since the sizes of modern databases, application files, or multimedia files can be extremely large.

A. Statement of the Problem

This paper work intended to come up with a new algorithm that optimizes data compression using Forward Difference Technique on Huffman's algorithm to compress and decompress text messages.

B. Aim and Objectives of the Study

The aim of this paper is to evaluate the performance of Forward Difference on Huffman compression algorithm without loss of information.
Some of the specific objectives include:

i. To provide algorithm for coding and decoding using Forward Difference on Huffman Coding Technique
ii. To develop a software framework for the implementation of the proposed algorithm
iii. To measure the performance of Huffman algorithm against the Forward Difference on Huffman using Compression Ratio, Compression Factor and Saving Percentage of the algorithms.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

II THEORITICAL FRAME WORK

Initial work on lossless data compression, according to [7], began in the late 1940s with the development of Information Theory. Claude Shannon and Robert Fano, in 1949, devised a methodical way to assign codewords based on probabilities of blocks. An optimal method for doing this was then found by David Huffman in 1951. Early implementations were typically done in hardware, with specific choices of codewords being made as compromises between compression and error correction. In the mid-1970s, the idea emerged of dynamically updating codewords for Huffman encoding, based on the actual data encountered. And in the late 1970s, with online storage of text files becoming common, software compression programs began to be developed, almost all based on adaptive Huffman coding, [14].

Until 1980, [7] said, most general-compression schemes used statistical modeling. Later, in 1977, Abraham Lempel and Jacob Ziv published their groundbreaking LZ77 algorithm, the first algorithm to use a dictionary base to compress data. LZ77 used a dynamic dictionary oftentimes called a sliding window. The drawback to LZ77 is that it has a small window size. Around the middle 80s, subsequent work by Terry Welch, the so-called LZW algorithm rapidly became the method of choice for most general-purpose compression systems. It was used in programs such as PKZIP, as well as in hardware devices such as modems and UNIX machines. In 1988, a man named Phil Katz came up with PKARC after he has studied ARC's popularity and decided to improve it by writing the compression and decompression routines in assembly language. The format was again updated in 1993, when Katz released PKZIP 2.0.

[9] associated two standard metrics often employed in data compression, which are efficiency and effectiveness. Efficiency is the measure of a resource requirement. Generally, it is the speed or throughput of an algorithm. It can be measured in CPU time (sec), symbols per second (sym/sec), or another similar hybrid measure. Effectiveness is the amount of redundancy removed. It is commonly expressed as a compression ratio (%), or in bits per symbol (bps).

[8] stated that the code used by most computers for text files is known as ASCII (American Standard Code for Information Interchange). ASCII can depict uppercase and lowercase alphabetic characters, numerals, punctuation marks, and common symbols. Other commonly used codes include Unicode etc.

Decoding is the opposite process of encoding, that is, the conversion of an encoded format back into the original sequence of characters. Encoding and decoding are used in data communications, networking and storage [6].

Data compression has a wide range of application as stated by [4], especially in data storage and data transmission. Some of these applications include:

It is used in many archiving systems such as ARC and, Telecommunications such as voice mail and teleconferencing. Audio, video, graphical, streaming media, Internet telephony applications and textual information can all benefit from data compression [11].

[3] identified two types of data Compression techniques, namely:

i. Lossless Compression
ii. Lossy Compression

Lossless compression techniques, [9] said, as their name implies, involve no loss of information. If data have been losslessly compressed, the original data can be recovered exactly from the compressed data. Lossless compression, he added, is generally used for applications that cannot tolerate any difference between the original and reconstructed data, and a good example is in the text compression.

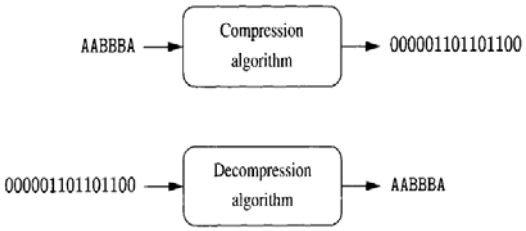


Figure2.1: Lossless Compression and Decompression Algorithm (source: Ida, 2006)

A lossy data compression according to [3] is lossy if it is not possible to reconstruct the original message exactly from the compressed version. He added that Lossy compression is called irreversible compression since it is impossible to recover the original data exactly by decompression. This means that there are some insignificant details that may get lost during the process of compression.

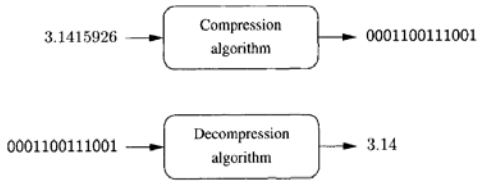


Figure2.2: Lossy Compression and Decompression Algorithm (source: Ida, 2006).

A. Shannon-Fano Algorithm

This is one of the earliest lossless compression techniques, invented in 1949 by Claude Shannon and Robert Fano. In the Shannon-Fano approach [3] said that a binary tree is constructed in a

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

'top-down' manner. This technique, according to [7], is to build a Shannon-Fano tree according to a specification designed to define an effective code table. The actual algorithm is simple:

The algorithm to generate Shannon-Fano codes is fairly simple

1: Parse the input, counting the occurrence of each symbol.
2: Determine the probability of each symbol using the symbol count.
3: Sort the symbols by probability, with the most probable first.
4: Generate leaf nodes for each symbol.
5: Divide the list in two while keeping the probability of the left branch roughly equal to those on the right branch.
6: Prepend 0 and 1 to the left and right nodes' codes, respectively.
7: Recursively apply steps 5 and 6 to the left and right subtrees until each node is a leaf in the tree.

Source: (Nelson & Gailly, 1996).

B. Huffman algorithm

Huffman Coding as lossless coding technique is another variant of entropy coding that works in a very similar manner to Shannon-Fano Coding, but the binary tree is built in a "top-down" manner in order to generate an optimal result [2]. The algorithm to generate Huffman codes:

1: Parse the input, counting the occurrence of each symbol.
2: Determine the probability of each symbol using the symbol count.
3: Sort the symbols by probability, with the most probable first.
4: Generate leaf nodes for each symbol, including P, and add them to a queue.
While (Nodes in Queue is greater than 1)
Remove the two lowest probability nodes from the queue.
Prepend 0 and 1 to the left and right nodes' codes, respectively.
Create a new node with value equal to the sum of the nodes' probability.
Assign the first node to the left branch and the second node to the right branch.
Add the node to the queue
5: The last node remaining in the queue is the root of the Huffman tree.

The working of Huffman's algorithm:
Suppose the following message is to be sent, as explained by[4]:
"XXXXXXXXXXXXXXXXXXXXYYYYYYYYYYYYYYYYYYYYZZZZZZZZZZZZ"

Table 1.2: Distinct Symbols and Frequencies from Huffman

Symbol	Z	Y	X	
Freq.	13	19	20	52

Source: Kattan, 2006

The algorithm follows as:

1: The number of distinct characters identified after removing redundancy and the count of their frequencies noted.
2: Create Probabilities table as in Table 2 above.
3: A binary Tree in Fig. 3 is then created bottom up, based on the probabilities, thereby giving those character with high frequency fewer bits to transmit with 1 on the right and 0 on the left.

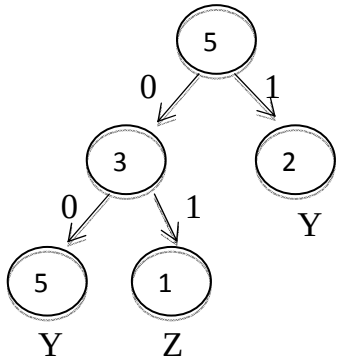


Figure 2.3: Resultant Huffman Tree from Table 2.2

Finally the bits streams to be transmitted are 00101, which are the codewords in the Table 2.3 below:

Table 2.2: Distinct Symbols and Huffman's codeword

Symbol	Z	Y	X
Codeword	00	01	1

Source: Kattan,2006

The average length of code transmitted is = (13*2)+(19*2)+(20*1) = 84 bits
Initially the number of bits to be transmitted without compression would have been = 52*8 = 416 bits. That means the amount of bits saved is (416 – 84) = 332 bits, which is about 80% of the original information or message.

Channel Probability is used in Information Theory to measure the outcome of independent

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

transmissions in a memoryless communication channel. [5] explained that if a word x is received, then for any codeword c ∈ C the forward channel probability is given by:

$$P(x\ received \mid c\ sent) = P^e(1-P)^{n-e} \quad (2)$$

where n is the length of x and e is the number of places at which x and c differ.

The knowledge of channel probability can be extended to data compression to improve the compressibility of a message coded using Fixed Length Codes (FLC). Equation (2) can be written as:

$$P_{new} = P^e(1-P)^{n-e} \quad (3)$$

where P_{new} is the new probability of a distinct symbol, P is the probability of 1s in the FLC, (1-P) the probability of 0s, length of e is the number of 1s in a codeword and (n-e) the number of 0s in the same codeword.

The forward difference is a finite difference scheme defined by:

$$\Delta a_n \equiv a_{n+1} - a_n \quad (4)$$

Where n = 0,1,2…

[12] said that in science, engineering, and mathematics, the Greek letter *delta* (Δ) is used to denote the *change* in a variable. The term *delta encoding* refers to several techniques that store data as the *difference* between successive samples (as residuals), rather than directly storing the samples themselves (as source symbols).

Delta encoding can be used for data compression when the values in the original data are *smooth*, that is, there is typically only a small change between adjacent values. Steven further explained that delta encoding has increased the probability that each sample's value will be near zero, and decreased the probability that it will be far from zero. This uneven probability is just the thing that Huffman encoding needs to operate.

III. METHODOLOGY

The proposed algorithm consists of the following:

A. Encoding Algorithm

1: Read the input file
2: Serialize each character
3: Determine the probability of each character
4: Determine the positions of each character
5: Compute the Forward Difference of the character positions using (4)
6: Transform the difference to twos complement code
7: Compute the new probability using the twos complement code using (3)
8: Build a Huffman's Binary Tree using the new probability and the distinct characters.
9: Determine the codeword for each distinct character by traversing the tree from root node to leaf node.
10: Determine the binary symbols to be transmitted for the whole message.

B. Decoding Algorithm

1: Read the whole coded message bit-by-bit
2: Determine a codeword from the coded message
3: Traverse a Huffman Tree from the root to the leaf node
4: Determine its symbol the code represents and its new probability
5: Use the new probability to reconstruct the twos complement code
6: Convert the twos complement code to its decimal equivalents
7: Determine which of the decimal numbers describes a delta difference
8: Use the delta differences and Backward Summation to regenerate the positions of each character
9: Use the character positions to allocate each character its positions
10: Write the whole message to a file.

A. Advantages of the System

The system compresses and decompresses information losslessly better than Huffman coding techniques.

IV. RESULTS AND DISCUSSION

The Lossless algorithm was tested on File1, File2 … File4 are text files of sizes 1224 bytes, 2548 bytes, 4096 bytes, and 8192 bytes with each having different contents.

The results in Table 4.1 and Table 4.2 were obtained from the implementation of "Data Compression using Forward Difference Technique on Huffman Algorithm", which are.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Table4.1: Compression using Huffman algorithm

Filename	Original File size(byte)	Compressed File size(byte)	Compression ratio	Compression factor	Saving Percentage
File1	1224	683	66.70	1.50	33.30
File2	2548	1252	61.13	1.64	38.87
File3	4096	2765	67.50	1.48	32.50
File4	8192	5025	61.34	1.63	38.66

Table4.2: Compression using Forward Difference on Huffman Algorithm (FDHA)

Filename	Original File size(byte)	Compressed File size(byte)	Compression ratio	Compression factor	Saving Percentage
File1	1224	631	61.62	1.62	38.38
File2	2548	1205	58.84	1.70	41.16
File3	4096	2567	62.67	1.60	37.33
File4	8192	4905	59.88	1.67	40.12

Results shown in Table4.1 and Table4.2 indicate that File1 of size 1224 bytes when compressed using Huffman algorithm yields a compression factor of 1.50 and 33.30 as the Saving percentage; while using FDHA has compression factor of 1.62 and a Saving advantage of 33.38. File2 of size 2548 bytes when compressed using Huffman algorithm yields a compression factor of 1.64 and a Saving advantage of 38.87; while using FDHA has compression factor of 1.70 and a Saving advantage of 41.16. File3 of size 4096 bytes when compressed using Huffman algorithm yields a compression factor of 1.48 and a Saving advantage of 32.50; while using FDHA has compression factor of 1.60 and a Saving advantage of 37.33. File4 of size 8192 bytes when compressed using Huffman algorithm yields a compression factor of 1.63 and a Saving advantage of 38.66; while using FDHA has compression factor of 1.67 and a Saving advantage of 40.12.

V. CONCLUSION

From the above discussion, it implied that the performance of FDHA was better than Huffman alone.

REFERENCES

[1] Choure, T. (2004), *Information Technology Industry in India*. Retrieved November 22, 2012 from http://books.google.com.ng/books

[2] Guy, B. (2010), Introduction to Data Compression. Retrieved October 3, 2012 from http://www.cs.cmu.edu/afs/cs/project/pscico-guyb/realworld/www/compression.pdf

[3] Ida, M.P. (2006), *Fundamental Data Compression*. Elsevier Publishers: New York. Retrieved April 23, 2013 from http://www.dbebooks.com

[4] Kattan, A. J. (2006), Universal Lossless Compression Technique with built in Encryption. Retrieved October 24, 2012 from www.ahmedkattan.com/Master_Dissertation.pdf

[5] Ling, S. (2004), *Coding Theory-A First Course*, Cambridge University Press: New York.

[6] Margaret, R. (2005), Encoding and Decoding. Retrieved January 23, 2013 from http://www.searchnetworking.techtarget.com/definition/encoding-and-decoding

[7] Nelson, M. & Gailly, J. (1996). The Data Compression Book 2nd Ed. Penguin Publishers: New York.

[8] Salomon, D. (2007), Data Compression - *The Complete Reference*, Springer-Verlag: London.

[9] Sayood, K. (2006), Introduction to Data Compression . Retrieved April 23, 2013 from http://www.csd.uoc.gr/~hy438/lectures/Sayood-DataCompression.pdf

[10] Sebastian, D. (2003), Universal lossless data compression algorithms. Retrieved August 31, 2012 from http://www.f5.com/pdf/white-papers/adv-compression-wp.pdf

[11] Shane, J.C. (2007), Efficient Data Representations for Information Retrieved January 27, 2013 from http://www.cs.rmit.edu.au/publications/sc07-thesis_final.pdf

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[12] Steven, S.W (2011) The Scientist and Engineer's Guide to Digital Signal Processing. Retrieved April 23, 2013 from www.dspguide.com/ch27/4.htm

[13] Twari, P. (2007), *Information Technology and Library Evolution*. Retrieved November 11, 2012 from http://books.google.com.ng/books

[14] Wolfram, S. (2002), History of Data Compression. Retrieved December 6, 2012 from http://www.wolframscience.com/compr.htm

Authors Profile

Adamu Garba Mubi is a postgraduate student from Adamawa State University, Mubi Nigeria. He obtained his B.Tech in Computer science from Abubakar Tafawa Balewa University Bauchi, Nigeria in 2004. He is currently an M.Sc student at Adamawa University. His main area of research is Data Compression. He is a Registered Member of

Dr. P. B. Zirra is a lecturer with Adamawa state University, Mubi Nigeria. He obtained his Doctorate Degree in Computer Science from Modibbo Adamawa University Yola in 2012, M.sc in Computer science from A.T.B.U Bauchi in 2006, MBA (Finance) from University of Maiduguri in 2000 and had his B.Tech. Computer 1994 from ATBU. His Area of interest include Computer Network and Security, he is happily married with two Children.

Wireless Sensor Networks Attacks and Solutions

Naser Alajmi
Computer Science and Engineering Department, University of Bridgeport
Bridgeport, CT 06604, USA

***Abstract*—A few years ago, wireless sensor networks (WSNs) used by only military. Now, we have seen many of organizations use WSNs for some purposes such as weather, pollution, traffic control, and healthcare. Security is becoming on these days a major concern for wireless sensor network. In this paper I focus on the security types of attacks and their detection. This paper anatomizes the security requirements and security attacks in wireless sensor networks. Also, indicate to the benchmarks for the security in WSNs**

Keywords-Wireless sensor network, security, vulnerability, attacks

I. INTRODUCTION

The security of wireless sensor networks is the area that has been discussed extremely through a few years ago. Networks have different applications. These applications comprise several levels of monitoring, tracking, and controlling. Wireless sensor networks consist of enormous number of small nodes. These nodes are deployed in some important areas. There are a group of applications that used for some purposes. So, in military application, sensor nodes include monitoring, battlefield surveillance and object tracking. The medical application, sensors can be helpful in patient diagnosis and monitoring. Most of these applications are deployed to monitor an area and to have a reaction when they record a sensitive factor [7]. Wireless sensor networks are emerging as both central new stage in the IT ecosystem and a rich area of active research involving hardware and system design networking, distributed algorithms, programming models, data management, security and social factors [2], [3]. Wireless sensor networks are vastly used in the area that is going to check for a particular task. Sensor nodes are liable to physical capture. Because the target of sensor nodes is low cost, tamper-resistant hardware is unlikely to take over.

The main goal of this paper is to find an energy efficient security solution thus to keep WSNs secured from any type of attacks. Also, this paper suggests several resolutions for the wireless sensor networks. Section II is analysis security requirements. Section III explains different security threats. Section IV makes a picture of benchmarks for WSN. Lastly, conclude on future works.

II. SECURITY REQUIREMENTS

Security in wireless sensor networks has to be comprehensives a fundamental of requirements. These requirements are not only guarantee safeguard of sensitive data but also to achieve bounded resources in each sensor node, which remains the sensor network alive. Attacker motivation and vulnerabilities, and opportunities are two factors, which give the attacker possibility impact to the wireless sensor networks.

A. Data Confidentiality

Data confidentiality is preserving the information hideaway from adversaries. The great way to keep the data invisible is to encrypt data with a secret key [1], [4]. The authorized can import data.

B. Data Authentication

The basis for several applications in wireless sensor network is message authentication. Data authentication blocks any part that illegal from engage in the network alongside original nodes must be eligible to reveal from unauthorized nodes [1]. Also, data is important to make sure that started from the accurate source and the node that is claimed must be the end of a connection.

C. Data Integrity

The adversaries have tried to change or modify the data. Therefore, data integrity makes sure the recipient who received message has not modified by unauthorized through transmission.

D. Data Freshness

Data freshness means that the data is a fresh. So it guarantees that no old data or messages have been replayed. Although data integrity and confidentiality are assured, there is a need to make sure the freshness for each message. Furthermore, the malicious node does not reply or resend old or previous data.

E. Access Control

Access control prevents unauthorized access to a resource. It should be able to prevent any participate in the network that is unauthorized.

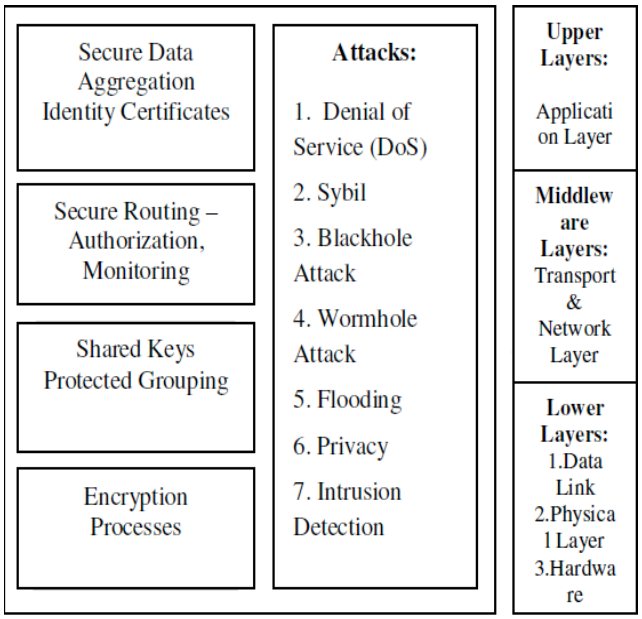


Figure 1. Figure 1: Security at various levels in Sensor Network

III. SECURITY ATTACKS ON WSN

Wireless sensor networks are very weak and susceptible to many types of security attacks cause to the broadcast. Also, the other reason is put the sensor nodes in a dangerous environment whether in public area or battlefield. The security threats and attacks in wireless sensor networks as follows:

A. Sybil Attack:

Wireless sensor network is vulnerable to the Sybil attack. In such a case, a node can be more than one node using different identities of legal nodes. Therefore, a single node presents multi identities to other nodes in the network [6], (Figure 2). Sybil attack tries to degrade the integrity of data, security and resource utilization that the distributed algorithm attempts to achieve [2]. Authentication and encryption mechanisms can prevent an outsider to launch a Sybil attack on the wireless sensor network. Public key cryptography can avoid such an insider attack, but it is too costly to be used in the resource constructed sensor networks [4]. Identities must be verified so Karlof and Wagner [5] said that, might be done using public key cryptography, but generating and verifying digital signatures is beyond the capabilities of sensor nodes. Newsome, Shi, Song, and Perrig [13] indicated to several defenses against Sybil attack by using radio resource testing, verification of key sets for random key predistribution, registration and position verification in sensor network. The probability of Sybil node being detect is:

$$\begin{aligned} Pr(detection) &= 1 - Pr(nondetection)^{r}_{1round} \\ &= 1 - (1 - Pr(detection)_{1round})^{r} \\ &= 1 - \left(1 - \sum_{allS,M,G} \frac{\binom{s}{S}\binom{m}{M}\binom{g}{G}}{\binom{n}{c}} \frac{S-(m-M)}{c}\right)^{r} \end{aligned}$$

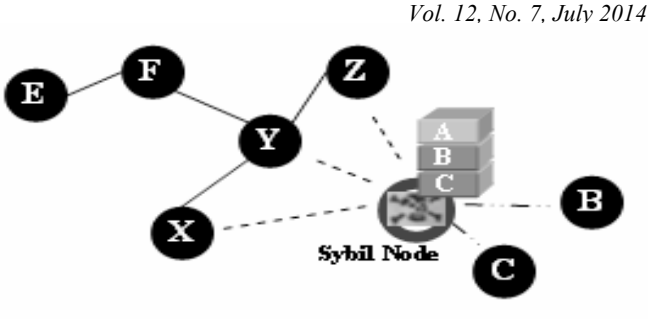


Figure 2. Figure 2: Sybil Attack

B. Wormhole Attack:

Wormhole attack is a significant attack in which the attacker records the packet at single location in the network and tunnels those to another location. The transmitting of bits could be done selectively. Wormhole attack is an important threat to wireless sensor network, because this type of attack does not need compromising a sensor in the network or other. It could be implemented at the initial phase when the sensor launch to discover the information. The wormhole attack is showed in Figure (3). Two malicious nodes X1 and X2, connected by a powerful connection, create a wormhole. Node A and node B select the shortest route provided by the wormhole for send data. Data will be caught by the malicious nodes and then by the attacker [7].

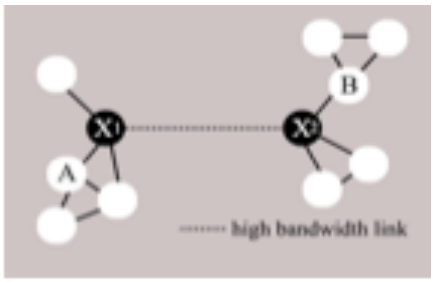


Figure 3. Figure 3: Wormhole Attack

C. Denial of Service Attack:

Denial of service attacks can disrupt wireless transmission and occur either unintentionally in the form of interference, noise or collision at the receiver side or at the context of attacks [8]. There are some targets that attackers need to reach them such as network access, network infrastructure, and server application. DoS attack attempt to drain the resources available to the victim node by transferring extra needless data. Therefore, prevents users to accessing services. Denial of Service attack is meant not only for the adversary's seek to subvert, disrupt, or destroy a network but also for any event that diminishes a network's capability to provide a service [10]. Denial of Service attacks is created in different layers. In the physical layer the DoS attack could be jamming and tampering, at link layer, collision, exhaustion, unfairness, at

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

network layer, neglect and greed, homing, misdirection, black hole and at transport layer DoS attack could be executed via malicious flooding and desynchronization [9].

The technique to prevent DoS attacks includes payment for network resources, pushback, strong authentication and identification of traffic [10]. There are some techniques to secure the reprogramming process thus one of them uses authentication flows. The choice for Denial of Service is the rekeying request packet. Hence it comes from the node only when any two consecutive keys are invalidated or lifetime of the keys has been expired. Therefore, if the rate of rekeying requests is coming frequently, then base station can conclude for possible DoS attack and drop the packet from the node for a configurable period of time [12]. The attacker find the re-keying request packet is a chance to send it again and make the DoS begins.

D. HELLO Flooding Attack:

This type of attack uses HELLO packet as a weapon to encourage the sensor in networks. The attacker with a high radio transmission range and processing power sends HELLO packet to a number of sensor nodes thus they are separated in a large area within a wireless sensor network [5]. The sensors are thus impacted that the adversary is their neighbor. There are several agreements require HELLO packet radio node to node nearby to its own broadcasting. Attacker with power to begin track broadcast, so that the network each node is believed to attack its neighbors [11]. A malicious node with a powerful connection, which transfer HELLO messages to nodes that the malicious node is a neighbor and will send data to it [5], (Figure 3).

E. Sinkhole Attack:

The sinkhole attack is an especially dangerous attack that prohibits the base station to gain entire and correct sensing data, consequently making a severe threat to the higher layer application (Figure 4). In a sinkhole, the adversary's goal is to lure nearly all the traffic from a particular area through a compromised node, creating a metaphorical sinkhole with the adversary at the center [5]. A compromised sensor node attempts to impact the information to it from any neighboring node. Thus, sensor node eavesdrops on each information is being communicated with its neighboring sensor nodes. Sinkhole attack works by making a compromised node look especially attractive to surrounding nodes with respect to the routing algorithm. For example, an adversary could spoof or reply an advertisement for an extremely high quality rout to the base station [5].

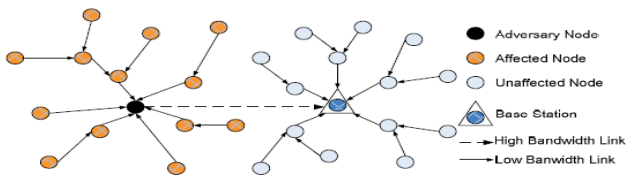


Figure 4. Figure 4: Sinkhole Attack

IV. SECURITY BENCHMARKS FOR WSN

In recent years, wireless sensor networks have grown and used for broad range of many applications such as weather, military aim tracking, and patient monitoring. Therefore, these sensor networks need protection from illegal attackers. There are some security benchmarks that sensor networks should be had.

A. Encryption

In fact, most of wireless sensor network hold in an open area or dangers location, thus it susceptible to the network attacks. Eavesdropping or add messages into the network are significant to WSN [14]. It has to take over key methods that protect WSN such as message authentication codes, symmetric key encryption and public key cryptography [1].

B. Data Partitioning

The technique of partitioning is to separate the data in networks into some or several parts. In wireless sensor networks, Deng J. [15] gives a solution to make sure the attacker cannot catch the information by using the data partitioning. Divide the data into multi packets so each packet transfers on a different route to nodes. At this point, the attacker tries to get all packets of a data from the network, thus it has to be capable to the entire networks. It is a perfect solution, but the energy consumption increased more than normal [7].

C. Secure Data Aggregation

Transmit data in wireless sensor network increased than before. As a result, the most issue in network is data traffic. So, the cost is rising. To reduce the high cost and network traffic, wireless sensor node aggregates measurements before transferring to the base station [1]. Wireless sensor network architecture, aggregation carries out in many locations in the network. Aggregation locations should be secured [16].

D. Cryptography

Symmetric key cryptography is a key that used in cryptography solutions in wireless sensor networks. Symmetric key is suitable and rapid to implement [7]. A cryptography method is used to prohibit some of the security attacks.

E. Shared Keys

A better deal of the concentration in wireless sensor networks is the field of key management. WSN is a single in this feature because size, mobility and power constraints, [1]. There are four types of keys management, global key, pair wise key node, pair wise key group, and individual key. Each one of these keys is a solution to prevent attacks in wireless sensor networks.

V. CONCLUSION

Security issue in wireless Sensor Network is more important than other issue. In recent years, security in WSN has frequently concerns. Wireless sensor networks are growing used in environment, commercial, health and military

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

applications. This paper briefs a sort of requirements that wireless sensor network have to be include and also introduce some of the security attacks. In addition wireless sensor network benchmarks.

REFERENCES

[1] Abhishek Jain, Kamal Kant, and M. R. Tripathy, "Security Solutions for Wireless Sensor Networks", to appear in IEEE ICACCT 2012.

[2] Al-Sakib Khan Pathan, Hyung-Woo Lee, Choong Sean Hong, "Security in Wireless Sensor Networks: Issues and Challenges", Proc. ICACT 2006, Volume 1, 20-22 Feb, 2006, pp. 1043-1048

[3] Culler, D. E and Hong, W., "Wireless Sensor Networks", Communication of the ACM, Val. 47, No. 6, June 2004, pp. 30-33

[4] Mayank Saraogi, "Security in Wireless Sensor Networks", University of Tennessee, Knoxville.

[5] Karlof, C. and Wagner, D., "Secure routing in wireless sensor networks: Attacks and countermeasures", Elsevier's Ad Hoc Network Journal, Special Issue on Sensor Network Applications and Protocols, September 2003, pp. 293-315.

[6] J. R. Douceur,(2002) "The Sybil Attack," in 1st International Workshop on Peer-to-Peer Systems(IPTPS"02).

[7] David Martins, and Herve Guyennet, "Wireless Sensor Network Attacks and Security Mechanisms: A Short Survey", 2010 IEEE.

[8] A. Perrig, J. Stankovic, and D. Wagner, "Security in Wireless Sensor Networks", Communications of the ACM, 47(6):53– 57, June 2004.

[9] Kalpana Sharma. M K Ghose, "Wireless Sensor Networks: An Overview on its Security Threats", IJCA Special Issue on Mobile Ad-hoc Networks 2010.

[10] David R. Raymond and Scott F. Midkiff,(2008) "Denial-of-Service in Wireless Sensor Networks: Attacks and Defenses," IEEE Pervasive Computing, val. 7, no. 1, 2008, pp. 74-81.

[11] Yan-Xiao Li, Lian-Qin and Qian-Liang, "Research on Wireless Sensor Network Security", In Proceedings of the International Conference on Computing and Security, 2010 IEEE.

[12] V. Thiruppathy Kesavan and S. Radhakrishnan, "Secret Key Cryptography Based Security Approach for Wireless Sensor Networks", International Conference on Recent Advances in Computing and Software Systems, 2012 IEEE.

[13] Newsome, J., Shi, E., Song, D, and Perrig, A, "The Sybil attack in sensor networks: analysis & defenses", Proc. of the third international Symposium on Information processing in sensor networks, ACM, 2004, pp. 259 – 268.

[14] Kalpana Sharma. M K Ghose, Deepak Kumar, Raja Peeyush Kumar Singh, Vikas Kumar Pandey, "A comparative Study of Various Security Approaches Used in Wireless Sensor Networks", In IJAST, Vol 7, April 2010.

[15] Deng J, Han R, and Mishra S. "Countermeasures against Traffic Analysis Attacks in Wireless Sensor Networks", IEEE, 2005, pp. 113-126.

[16] Pathan, A. S. K., Hyung-Woo Lee, and Choong Seon Hong "Security in Wireless Sensor Networks: Issues and Challenges" Advanced Communication Technology (ICACT), 2006.

AUTHORS PROFILE

Naser Alajmi is a PhD student with the Computer Science and Engineering Department, University of Bridgeport, Bridgeport, CT, USA (e-mail: nalajmi@my.bridgeport.edu)

Enhancing the Accuracy of Biometric Feature Extraction Fusion Using Gabor Filter and Mahalanobis Distance Algorithm

[1]Ayodeji S. Makinde, [2]Yaw Nkansah-Gyekye, [3]Loserian S. Laizer

[1,2,3] School of Computational and Communication Science and Engineering, NM-AIST, Tanzania

***Abstract*- Biometric recognition systems have advanced significantly in the last decade and their use in specific applications will increase in the near future. The ability to conduct meaningful comparisons and assessments will be crucial to successful deployment and increasing biometric adoption. The best modality used as unimodal biometric systems are unable to fully address the problem of higher recognition rate. Multimodal biometric systems are able to mitigate some of the limitations encountered in unimodal biometric systems, such as non-universality, distinctiveness, non-acceptability, noisy sensor data, spoof attacks, and performance.**

More reliable recognition accuracy and performance are achievable as different modalities were being combined together and different algorithms or techniques were being used. The work presented in this paper focuses on a bimodal biometric system using face and fingerprint. An image enhancement technique (histogram equalization) is used to enhance the face and fingerprint images. Salient features of the face and fingerprint were extracted using the Gabor filter technique. A dimensionality reduction technique was carried out on both images extracted features using a principal component analysis technique. A feature level fusion algorithm (Mahalanobis distance technique) is used to combine each unimodal feature together. The performance of the proposed approach is validated and is effective.

Keywords – Gabor filters; Mahalanobis distance; principal component analysis; face; fingerprint; feature extraction.

I. INTRODUCTION

A. Background

With the advancement in networking, communication, and mobility in today's electronically wired information society, the need for accurate and reliable feature extraction of biometric traits in multibiometric systems is very crucial. Feature extraction refers to the process of generating a compact but expressive digital representation of the underlying biometric trait, called a template which contains only the salient discriminatory information that is essential for recognizing the person [2][4][10][24].

A good biometric is characterized by the use of features that are highly unique, so that the chance of any two people having the same characteristics will be minimal, stable, does not change over time, easily captured in order to provide convenience to the user, and prevent misrepresentation of the feature [19]. For multi-biometric recognition to have low False Rejection Rate (FRR) and False Acceptance Rate (FAR), an efficient feature extraction algorithm is needed.

In order to provide accurate recognition of individuals, the most discriminating information present in a bimodal face and fingerprint system must be extracted so that comparison between templates can be made. In this paper two prominent modalities (face and fingerprint images) are used for the bimodal features extraction. Gabor filter is used in extracting the salient feature from the two modalities, due to its ability to extract maximum information from local image regions, and being invariant against translation, rotation, and variations [1][7][8]. Although there are other algorithms which also perform better with different characteristics in feature extraction of face and fingerprint images, each of these algorithms cannot be used in extracting both face and fingerprint features at the same time. [6] used Gabor filter for the extraction of fingerprint features with an accuracy of 97.2%. Likewise, [13] also used Gabor filter for fingerprint feature extraction.

In the recent years Gabor filter was noted for the extraction of face images [1][28] than that of the fingerprint images. The problem in multimodal biometric fusion is that each modality was extracted with different algorithm which makes it difficult to fuse together [2][3]. However, based on empirical literature, it was noted that Gabor filter has a good feature extraction accuracy on both images especially at the feature extraction level.

B. Biometric system

Biometric systems are technologies used to identify or recognize individuals based on their physiological or behavioral characteristics. They capture biometric features of a person and extract a set of salient features that are compared with a set of features in the form of a template already extracted and stored in the database of the same person. Several human characteristics

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

that can be used as the basis for biometric systems include a person's face, fingerprint, iris, DNA. A biometric-based authentication system consists of two main phases, namely, enrollment and recognition [2][3].

C. Why multimodal biometric system?

Unimodal biometric systems are affected by the following problems [15][16]:

- Noisy sensor data: The sensed data might be noisy or distorted.
- Non-universality: While every user is expected to possess the biometric trait being acquired, in reality, it is possible for a subset of the users to be unable to provide a particular biometric.
- Lack of individuality: While a biometric trait is expected to vary significantly across individuals, there may be large similarities in the feature sets used to represent these traits. This limitation restricts the discriminality provided by the biometric trait.
- Spoof attacks: An impostor may attempt to spoof the biometric trait of a legitimate enrolled user in order to circumvent the system.

Due to these practical problems, the error rates associated with unimodal biometric systems are quite high which make them unacceptable for deployment in critical security applications. Some of the problems that affect unimodal biometric systems can be overcome by using multimodal biometric systems. Fusing two or more traits together tends to address some of the problems being faced in the unimodal biometric systems. In this paper, face and fingerprint images were fused together for biometric recognition due to the fact that they have some advantages over other modalities such as availability, collectability, mostly country or religion conflict free and long existence. Although there are other modalities such as iris, retina, and vein pattern with high uniqueness, universality, permanence, performance, resistance to circumvention, they are difficult to maintain. Thus, it makes face and fingerprint modalities to be some of the most researched and mature fields of authentication. Figure 1 illustrates different fusion of biometrics and various fusion scenarios that are feasible in multimodal biometric systems.

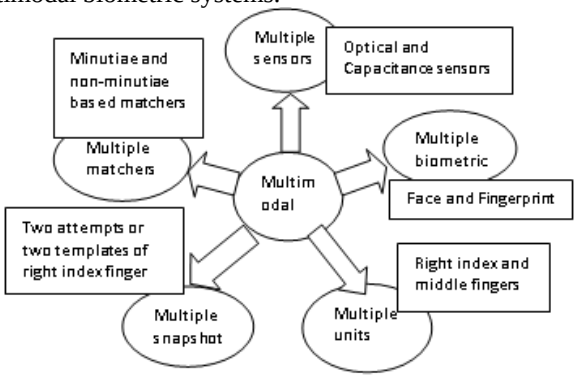


Figure 1: Scenarios in a multimodal biometric system [20]

This paper proposes an enhanced biometric feature extraction fusion using Gabor filter and Mahalanobis distance. The next section presents the literature review. Section 3 presents the material and method used. Results and discussion are given in Section 4. Conclusion and recommendations are made in the last section.

D. Originality and contribution

In this paper, we present a novel way of extracting the features of face and fingerprint modalities which form a bimodal biometric system using a Gabor filter. The extraction rate obtained from the proposed method was compared with other forms of feature extraction algorithms used in both face and fingerprint. The contribution of this paper was to enhance the accuracy and performance of multimodal biometric systems at the feature extraction level. One of the challenges of multimodal biometric systems is the difficulty of fusing two modalities of different feature extraction algorithm at the feature extraction level.

II. Literature review

In this section, we review the different biometric feature extraction algorithms that have been used for both face and fingerprint recognition.

A. Face feature extraction algorithm review

[12] presented an automatic facial feature extraction method using genetic algorithm. The method was based on the edge density distribution of the image. In the preprocessing stage a face is approximated to an ellipse and a genetic algorithm is applied to search for the best ellipse region match in which genetic algorithm is applied to extract the facial features. The experimental results validate that the proposed method is capable of automatically extracting features from various video images effectively under natural lighting environments and in the presence of a certain amount of artificial noise and of multiface oriented with angles. They proposed that the algorithm can be improved so that it can be applied to real world problem, by incorporating more characteristics in the fitness function during the evolutionary process and also to extract features on faces containing either beard or mustache.

[1] presented a neural network system for face recognition in which Gabor filter was used for extracting the salient features, and the feature vectors which were based on Fourier Gabor filters is used as input of their classifier which is a Back Propagation Neural Network (BPNN). Also, due to the large dimension of the input vector, a dimensionality reduction algorithm called Random Projection (RP) was used. Their experimental results demonstrate that using Gabor filter for face feature extraction proves the robustness of their solution, due to the fact that more salient features were being extracted.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[20] used independent component analysis (ICA) method for the feature extraction of face images for face recognition. They used a version of ICA derived from the principle of optimal information transfer through sigmoidal neurons. ICA was performed on face images in the FERET database under two different architectures, one of which treated the images as random variables and the pixels as outcomes, and a second one treated the pixels as random variables and the images as outcomes. The first architecture found spatially local basis images for the faces. The second architecture produced a factorial face code. The experimental results express that both ICA representatives were superior to representation based on PCA for recognizing faces across days and changes in expression. Also, a classifier that combined the two ICA representations gave the best performance.

[22] presented in their paper, a new model bidirectional associative memory (BAM) inspired architecture that can ultimately create its own set of perceptual features. The resulting model inherits properties such as attractor-like behavior and successful processing of noisy inputs, while being able to achieve principal component analysis (PCA) tasks such as feature extraction and dimensionality reduction. The model is tested by simulating image reconstruction and blind source separation tasks. Simulations show that the model fares particularly well compared to existing neural PCA and independent component analysis (ICA) algorithms. It is argued that the model possesses more cognitive explanative power than any other nonlinear/linear PCA and ICA algorithm.

[21] stated that LDA is one of the most commonly used techniques for data classification and dimensionality reductions. They further explain that LDA easily handles the situation where the within-class frequencies are unequal and their performance has been examined on randomly generated test data. This method maximizes the rate of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability. The prime difference between LDA and PCA is that the PCA does more of feature classification and LDA does data classification. In PCA, the shape and location of the original data set changes when transformed into a different space, whereas LDA does not change the location, but only tries to provide more class separability and draws a decision region between the given classes.

[24] proposed a weighted 2D Principal Component Analysis (2DPCA) model which addresses some of the challenges faced in using 2DPCA as a face recognition extractor. [27] make use of 2DPCA in the feature extraction of face recognition in which face images were represented in matrices or 2D images form compared to the conventional PCA which represents images in vector form. Also, they stated that using 2D images directly is quite simple and local information of the original images is preserved appropriately, which may bring more important features for facial representation. With all these advantages, not all the face images are easy to recognize. For example, frontal face images are easier to be recognized than profile face images, which subsequently led to the proposed weighted-2DPCA model to deal with some practical situations in which some face images in database are difficult due to their poses (front of the profile) or their qualities (noise, blur).

The algorithm used in a Weighted-2DPCA model for the face model construction consists of the following steps:

Step 1: Compute the mean image

$$\tilde{A} = \frac{\sum_{i=1}^{N} w_i A^{(i)}}{\sum_{i=1}^{N} w_i} \quad (1)$$

Step 2: Compute the matrix

$$G = \frac{\sum_{i=1}^{N} w_i (A^{(i)} - \tilde{A})^T (A^{(i)} - \tilde{A})}{\sum_{i=1}^{N} w_i} \quad (2)$$

Step 3: Compute eigenvectors $\{\Omega_1, \Omega_{2,\ldots,} \Omega_n\}$ and eigenvalues $\{\lambda_1, \lambda_{2,\ldots,} \lambda_n\}$ of G, where w_i is the weight and A is the image matric.

The results indicate that 97.3% accuracy using Weighted-2DPCA compared to 96.2% accuracy using 2DPCA and 95.2% accuracy using conventional PCA.

B. Fingerprint feature extraction algorithm review

[6] presented a Gabor filter based method for directly extracting fingerprint features from gray level images without the introduction of preprocessing of the original images acquired at the sensor level. Their proposed method is more suitable than conventional methods for a small scale fingerprint recognition system. Their experimental results demonstrate that the recognition to the k-nearest neighbor classifier using the proposed Gabor filter based features has an accuracy of 97.2% with 3-NN classifiers.

[14] presented a fingerprint feature extraction method through which minutiae are extracted directly from original gray-level fingerprint images without binarization and thinning. Their algorithm improves the performance of the existing ones along this stream by using the following steps: first, they preprocess the fingerprint images by making use of Gabor filter. Second, they find the computation of the orientation field which is very crucial in order to trace the ridgelines in the fingerprint images correctly. Third, they determined the starting pixels to trace ridgelines, once a starting pixel in a ridge has been decided, the minutiae are recorded. Their experimental results indicate that the approach can achieve better performance in both efficiency and robustness.

[9] presented a novel ridge tracing approach for extraction of fingerprint features directly from gray scale images. With this method, they made use of contextual information gathered

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

during the tracing process to better handle noisy regions. Their experimental results have been compared with other feature extraction algorithms such as Gabor based filtering as well as the original ridge tracing work, which clearly show that their proposed approach makes a ridge tracing more robust to noise and makes the extracted features more reliable.

[13] presented a set of fingerprint recognition algorithms which includes Gamma controller normalization and equalizer, fingerprint image division, fingerprint image binarization and different direction Gabor filter for feature extraction by taking into account both the global and local features of the fingerprints which were based on the fingerprint image enhancement and the texture using Gabor filter. The experimental results illustrate that the proposed algorithm can avoid all sorts of false characters more effectively and the recognition rate is higher than that of the traditional algorithm in the same conditions.

[26] made a comparative study involving four different feature extraction techniques for fingerprint classification. Also, they proposed a rank level fusion scheme for improving classification performance. They compared two well-known feature extraction methods based on Orientation Maps (OMs) and Gabor filters with two other new methods based on minutiae maps and orientation collinearity. Each feature extraction method was compared together in terms of accuracy and time. Moreover, they investigated on improving the classification performance using rank-level fusion. During the evaluation of each feature extraction method, their experimental results show that OMs performed best, in which Gabor feature fell behind OMs mainly because their computation is sensitive to errors in localizing the registration point. When fusing the rankings of different classifiers, they found that combinations involving OMs improve performance. Generally, the best classification results were obtained when they fused orientation map with orientation collinearity classifiers.

III. Proposed bimodal feature extraction fusion

The proposed idea gives rise to an innovative way to fuse the features of two different modalities/traits. In the case of this paper face and fingerprint images were used. The procedure of the proposed bimodal fusion is grouped in the following basic stages:

1. Preprocessing stage, which involves the use of histogram equalization in enhancing the image(s) to be recognized.
2. Features of each modality on the preprocessed image are extracted using Gabor filter.
3. Dimensionality reduction using principal component analysis to reduce the dimension of the feature vectors due to their high dimensionality.
4. The fusion stage combined the corresponding feature vectors of each modality. The features are fused using the Mahalanobis distance. These distances are normalized by applying hyperbolic and fused using the average sum rule tanh [25].

A. Histogram equalization

An image histogram is a graphical representation of the tonal distribution in a digital image [11]. When viewing an image represented by a histogram, what really happens is analyzing the number of pixels vertically with a certain frequency horizontally. In essence, an equalized image is represented by an equalized histogram where the number of pixels is spread evenly over the available frequencies. These respective areas of the image that first had little fluctuation will appear grainy and rigid, thereby revealing other unseen details.

In order to equalize the face and fingerprint image histogram the cumulative distribution function (cdf) has been computed. The cdf of each gray level is the sum of its recurrence and the recurrence of the previous gray level in the image. The histogram equalization equation is given as:

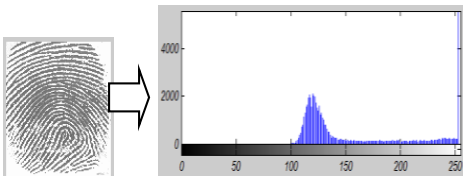
$$h(v) = round\left(\frac{cdf(v)-cdfmin}{(W\times H)-cdfmin} \times (N-1)\right) \quad (3)$$

where cdf_{min} is the minimum value of the cumulative distribution function, W and H are the width and height of image, N is the number of gray levels used. This result is an equalized and normalized image.

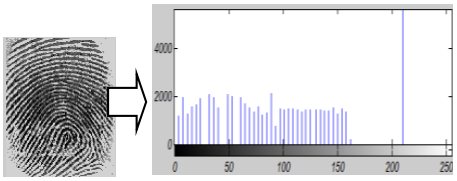
1. Histogram equalization for fingerprint image enhancement

Fingerprint images are not 100 percent perfect, they may be affected by noise due to some factors, including: noise on capturing devices, the tip of the finger from which the measurements are taken and ridge patterns may be affected by cuts, dirt, or wear and tear. Also, it was noted that most of the fingerprint feature extraction is carried out using minutiae-based method and image-based method, which at times require an extensive preprocessing operation which includes normalization, segmentation, orientation, ridge filtering, binarization, and thinning. With all these stages of preprocessing there is degradation in the image to be extracted, whereby reducing the salient features [5][25]. In our paper histogram equalization was adopted, which enhanced the contrast of the fingerprint image (increasing the quality of the image). Histogram equalization is capable of revealing unseen details in an image. The result of the enhanced contract image is shown in Figure 2. Which produced an enhanced fingerprint image that is suitable in feature extraction.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

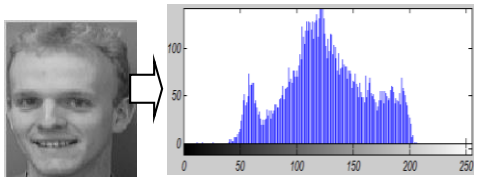


(a)

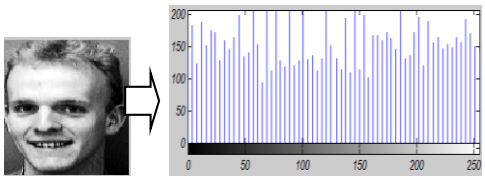


(b)

Figure 2: Fingerprint Image (a) before enhancement (b) After enhancement (histogram equalization)



(a)



(b)

Figure 3: Face Image (a) before enhancement (b) After enhancement (histogram equalization)

2. *Histogram equalization for face image enhancement*

Face is considered to be the most commonly used biometric trait of humans, since it has shown its importance over the last ten years or so. Not only is it an intensely researched area of image analysis, pattern recognition and even biometrics to be precise [18] but it has also become an important in our daily lives since it was introduced as one of the methods for identification to be used in e-passports [18]; we recognize each other and in many cases establish our identities based on faces. Some of the benefits of using face image as one of the bimodal traits in this paper are: it is not intrusive, can be done from a distance even without the users being aware they are being scanned, and can also be used for surveillance purposes (as in searching for wanted criminals, suspected terrorists, or missing children).

Face images are not 100 percent perfect as long as they can be affected by many factors such as environment condition (stress), age, pose, illumination, facial expression, as well as changes in appearance due to make-up, facial hair [2]. Some of these problems are solved through preprocessing of the face image. In this paper, histogram equalization was used in enhancing the image contrast, in order to reveal some of the unseen details in the face image. Figure 3 shows the result of the enhanced contract image produced which is suitable for feature extraction.

B. *Gabor filter*

Gabor filter is a linear filter which is used for edge detection. Its frequency and orientation are similar to that of the human visual system, and they have been found to be appropriate for texture discrimination and representation. Gabor filters are formed by modulating a complex sinusoid by a Gaussian function. Gabor filters have been used widely in pattern analysis application, and it has been proved in extracting more salient features both in the face [1] and fingerprint [6][14][26] images, which are the two modalities being used in this paper. A set of Gabor filters with different frequencies and orientations was used for extracting salient features from both face and fingerprint images. It is invariant against translation, rotation, and variations due to illumination and scale. Gabor filter also presents desirable characteristics of spatial locality and orientation selectivity. During feature extraction the dimension or size of the image does not change. For instance, in this paper the dimension of both face and fingerprint is 112x92, after applying Gabor filter to extract the salient features the dimension still remains the same. This motivated the use of the principal component analysis (PCA) as a dimensionality reduction technique.

Each filter of Gabor is defined by:

$$G_{\lambda,\theta,\varphi,\sigma,\gamma}(x,y) = \exp\left(-\frac{x'^2 - \gamma y'^2}{2\sigma^2}\right) cos\left(2\pi\frac{x^2}{\lambda} + \varphi\right)$$

where:

$$x' = xcos\theta + ysin\theta, y' = -xsin\theta + ycos\theta\ ,$$

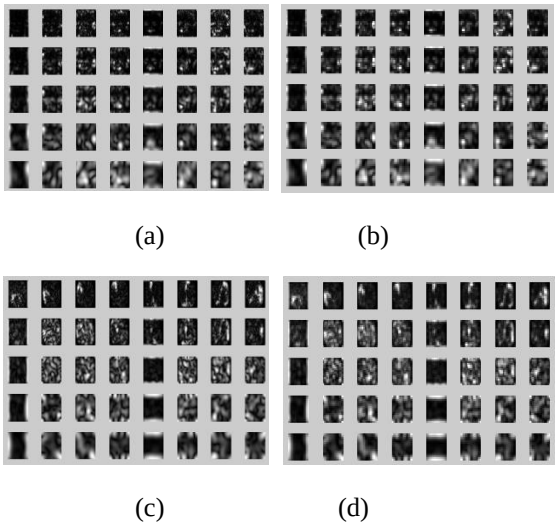
and λ , θ , φ, σ and γ are wavelength, orientation(s), phase offset(s), bandwidth and aspect ratio respectively.

1. *Gabor filter techniques for extracting face and fingerprint features*

In this paper, the enhanced face and fingerprint images are used to extract the identifiable features in face and fingerprint. Figure 4 shows the performance of the new method when the number of orientations varies and the number of scales was fixed. The feature extraction information was based on the use

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

of Gabor filtering parameters which includes: wavelength, orientation(s) degree, phase offset(s) aspect ratio, and bandwidth with their respective values.



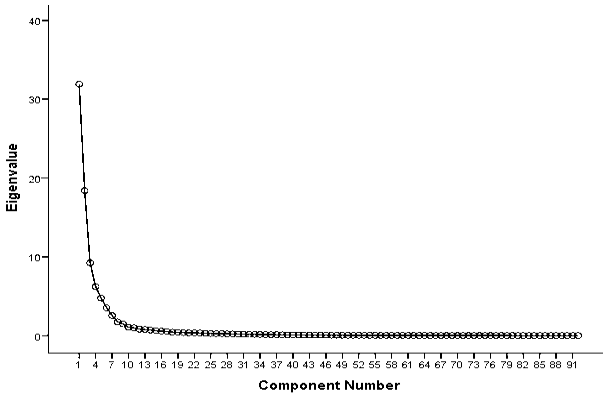
(a) (b)

(c) (d)

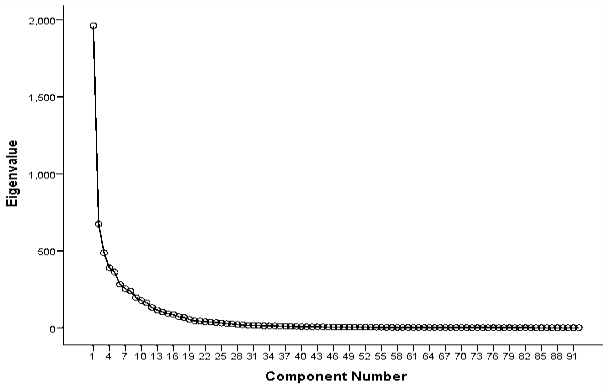
Figure 4: Gabor filter transformation of a sample of the face and fingerprint. (a) & (c) Magnitude responses of the filtering operation with the Gabor filter bank with no downsampling respectively (b) & (d) Magnitude responses of the filtering operation with the Gabor filter bank with downsampling 64 respectively.

C. Principal component analysis (PCA)

The curse of dimensionality is a major problem in the extracted feature vector during the extraction of the salient features on both face and fingerprint images. This is as a result of higher dimensional space, which results in an enormous amount of data to be required to learn. PCA is one of the most commonly used dimensionality reduction techniques. It finds the principal components which is the underlying structure in the data. Figure 5 shows the performance of the PCA on both face and fingerprint by finding their PC. It helps in speeding up the algorithms and reduces space used by the data during the training, validation and testing stages using a supervised learning neural network (MLP) and also improves how we display information in a tractable manner for human consumption.



(a)



(b)

Figure 5: Principal component performance on (a) face (b) fingerprint.

D. Mahalanobis distance technique

1. Fusion of face and fingerprint feature vectors

[25] proposed a novel technique mahalanobis distance technique in the fusion of fingerprint and iris feature vectors at the feature extraction level. They compared this technique with other techniques and it was proved to be easier and more effective to use, due to the facts that other feature fusion performed serially or parallel, which at the end results in a high dimensional vector. But their proposed algorithm generates the same size fused vector as that of unimodal.

In this paper, the feature vectors extracted from the input images are combined together to form a new feature vector by making use of the adopted method feature level fusion technique. The extracted features from face and fingerprint are homogeneous, each vector are processed to produce the fused vector of the same dimension.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

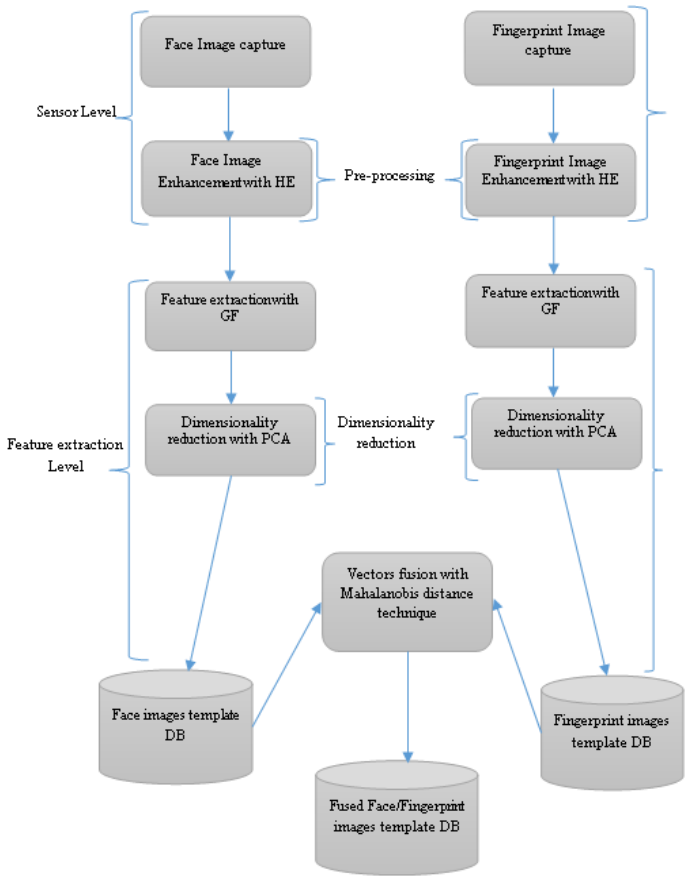


Figure 6: Architecture for the enhanced accuracy of biometric feature extraction fusion

IV. Experimental Results

The performance evaluation of the proposed method is analyzed using the ORL face database and the ATVS fingerprint database. The ORL face database (http://www.face-rec.org/databases/) contains images from 40 individuals, each providing 10 images of different pose, expressions and decorations. The ATVS fingerprint database (http://atvs.ii.uam.es/databases.jsp) contains images from 17 individuals, each providing 4 different positions each of right middle and the small finger. Based on the numbers of fingerprint per each individuals, the experiments are performed using the first four images samples per class for testing the new approach.

A. Comparison of face, fingerprint and the fusion

The results in Figure 7 show the performance rates of the extracted features for the face, fingerprint and the fusion of the two modalities. From the results, it was proven that the fusion of the two modalities have performance rate compared to the individual modalities.

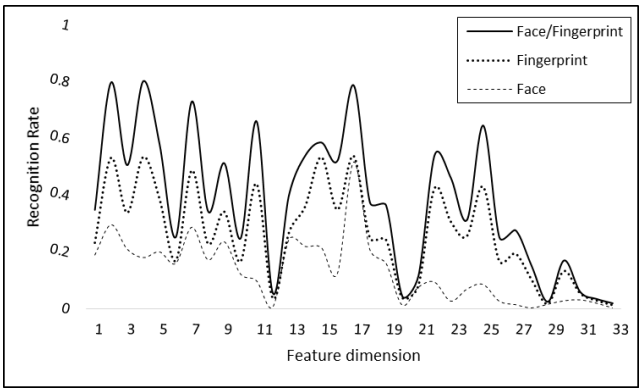


Figure 7: Recognition rate for different modalities

B. Comparison with other feature extraction methods

In this section the Gabor filter based feature extraction is compared to the KFA, LDA, PCA, and KPCA as shown in Table 1 for the face features, Table 2 for the comparison between Gabor filter and minutia based feature extraction for the fingerprint features, and Table 3 presents the results for the fusion of the face and fingerprint modalities using Gabor filter algorithm compared with other forms of feature extraction algorithms.

TABLE I: FEATURE EXTRACTION PERFORMANCE FOR FACE IMAGES

	KFA	**LDA**	**PCA**	**KPCA**	**Gabor filter**
Face	82.35 %	91.18 %	97.01 %	96.54 %	97.35 %

TABLE II: FEATURE EXTRACTION PERFORMANCE FOR FINGERPRINT IMAGES

	Minutia Based	**Gabor filter**
Fingerprint	97.76%	98.87%

TABLE III: FEATURE EXTRACTION PERFORMANCE FOR THE FACE AND FINGERPRINT FUSION.

	KFA/Minutia Based	**LDA/ Minutia Based**	**PCA/Minutia Based**	**KPCA/ Minutia Based**	**Gabor filter**
Face/fingerprint	90.01%	94.47%	97.39%	97.15%	98.11%

When compared with what other researchers have done, the proposed method of fusing face and fingerprint modalities

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

together using the Gabor filter technique for their feature extraction and Mahalanobis distance technique for fusing the two feature vectors together look promising.

V. Conclusion and recommendation

This paper presents an efficient way of extracting salient features in bimodal biometric systems. The proposed method uses face and fingerprint modalities from which the features were extracted. An image enhancement histogram equalization technique is used to enhance the face and fingerprint images. Salient features of the face and fingerprint were extracted using the Gabor filter technique. A dimensionality reduction technique is carried out on both images extracted features using a principal component analysis technique. A feature level fusion algorithm is used to combine each unimodal feature using the Mahalanobis distance technique. The performance of the proposed approach is validated and compared with other methods.

VI. Future work

In order to fully enhance the accuracy and performance of the biometric recognition system, the proposed method can be trained and tested using a multilayer perceptron neural network which is a powerful non-linear classifier, produces elegant solutions built of continuous basic functions, has the ability to handle noisy data, and is fast to run.

Acknowledgment

We sincerely appreciate the Nelson Mandela African Institution of Science and Technology through the School of Computational and Communication Science and Engineering for supporting this work to completion.

References

[1] A. Bouzalmat, N. Belghini, A. Zarghili, J. Kharroubi and A. Majda, “Face Recognition Using Neural Network Based Fourier Gabor Filters & Random Projection”. International Journal of Computer Science and Security (IJCSS), Vol. 5: Issue 3. 2011

[2] A. K. Jain, A. A. Ross, and K. Nandakumar, “Introduction to Biometrics”. Springer, 2011.

[3] A. K. Jain, A. Ross, and S. Prabhakar, “An introduction to Biometric Recognition,” IEEE Trans. on Circuits and Systems for Video Technology, vol. 14, pp. 4–20, Jan 2004.

[4] A. K. Jain, J. Feng, K. Nandakumar, "Fingerprint Matching", The IEEE Computer Society, 0018-9162/10, 2010

[5] A. Rattani, D. R. K. Ku, M. Bicege and M. Tistarelli, “Feature Level Fusion of Face and Fingerprint Biometrics”. Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on. IEEE, 2007.

[6] C. Lee and S. Wang, “Fingerprint Feature Extraction Using Gabor Filters”. Electronics Letters 35.4 (1999): 288-290.

[7] C. Theriault, N. Thome, M. Cord, “Extended Coding and Pooling in the HMAX Model”. Image Processing, IEEE Transactions on 22.2 (2013): 764-777.

[8] C. Yao and S. Chen, “Retrieval of Translated, Rotated and Scaled Color Textures”. Pattern Recognition 36.4 (2003): 913-929.

[9] D. Arpit and A. Namboodiri, “Fingerprint feature extraction from gray scale image by ridge tracing”. Biometrics (IJCB), 2011 International Joint Conference on. IEEE, 2011.

[10] D. Maltoni, D. Maio and A. K. Jain, “Handbook of Fingerprint Recognition”. Springer, 2009

[11] E. Sacco, J. R. M. i-Rubió, C. Regazzoni, and S. Maludrottu. "Image Preprocessing Methods for Face Recognition Applications."Univ. of. GENOVA (2010).

[12] G. G. Yen and N. Nithianandan, “Facial Feature Extraction using Genetic Algorithm”. Evolutionary Computation, 2002. CEC'02. Proceedings of the 2002 Congress on. Vol. 2. IEEE, 2002

[13] H. Jia, “Preprocessing and Feature Extraction, Coding, Matching Algorithm for Fingerprint Image”. Advanced Materials Research Vols. 805-806, pp. 1900-1906, 2013

[14] J. Yang, L. Liu, and T. Jiang, Member, IEEE, “An Improved Method for Extraction of Fingerprint Features”. Proc. the 2nd Int. Conf. Image and Graphics, Anhui, PR China. 2002

[15] Jain, Anil K., and Arun Ross. “Multibiometric systems.” Communications of the ACM 47.1 (2004): 34-40.

[16] Jain, Anil K., Arun Ross, and Sharath Pankanti. “Biometrics: a tool for information security.” Information Forensics and Security, IEEE Transactions on1.2 (2006): 125-143.

[17] K. Delac, M. Grgic, S. Grgic "Image compression effects in face recognition systems." Face Recognition (2007): 75-92.

[18] K. Delac, S. Grgic, and M. Grgic, “Image Compression in Face Recognition-a Literature Survey”. Recent Advances in Face Recognition, edited by: Kresimir Delac, Mislav Grgic and Marian Stewart Bartlett (2008): 236.

[19] L. Masek, “Recognition of Human Iris Patterns for Biometric Identification”, Master’s Thesis, The University of Western Australia, 2003

[20] M. S. Bartlett, J. R. Movellam, and T.J. Sejnowski, “Face Recognition by Independent Component Analysis”. IEEE Transactions on Neural Networks. 13(6): 1450-1464 2002

[21] S. Balakrishnama, A. Ganapathiraju, “Linear Discriminant Analysis - A Brief Tutorial”, Institute for Signal and Information Processing, 2010

[22] S. Chartier, G. Giguere, P. Renaud, J. Lina, and R. Proulx, “FEBAM: A Feature-Extracting Bidirectional Associative Memory”. Neural Networks, 2007. IJCNN 2007. International Joint Conference on. IEEE, 2007

[23] S. Prabhakar and A. K. Jain, “Decision-level Fusion in Fingerprint Verification,” Pattern Recognition, vol. 35, no. 4, pp. 861–874, 2002.

[24] T. H. Le, L. Bui, “Face Recognition Based on SVM and 2DPCA”, arXiv preprint arXiv:1110.5404 (2011).

[25] U. Gawanda, M. Zaveri, and A. Kapur, “A Novel Algorithm for Feature Level Fusion Using SVM Classifier for Multibiometrics-Based Person Identification”. Applied Computational Intelligence and Soft Computing 2013 (2013): 9.

[26] U. Rajanna, A. Erol, and G. Bebis, “A comparative study on feature extraction for fingerprint classification and performance improvements using rank-level fusion”. Pattern Analysis and Applications 13.3 (2010): 263-272.

[27] Y. Jian, L. Zhao, “Two-dimensional PCA: A New Approach to Appearance-based Face Representation and Recognition," Pattern Analysis and Machine Intelligence, IEEE Transactions, vol. 26, pp. 131-137, 2004.

[28] Y. Jin and Q. Ruan, “Face Recognition Using Gabor-Based Improved Supervised Locality Preserving Projections”. Computing and Informatics, Vol. 28, Pp. 81-95. 2012

GPGPU based Parallel Spam Filter

Prachi Goyal Juneja
M.Tech Scholar
Maulana Azad National Institute of Technology
Bhopal(M.P) India-462003

R.K.Pateriya
Associate Professor
Maulana Azad National Institute of Technology
Bhopal(M.P) India-462003

***Abstract*- Spam means unwanted emails in our mailboxes each day. These emails consist of promotional messages from companies, viruses, lucrative offers of earning extra income and many more. They are sent in bulk to flood our mailboxes and come from unknown sources. Various ways have been devised to deal with spam; these are known as Spam Filtering Techniques. Spam Filtering is done based on many parameters like keywords, URL, content etc. Content based spam filtering is becoming famous since it incorporates the judging of the email content and then analyzing it to be spam or ham. As the data is increasing and electronic data taking over most of the communication medium, one needs faster processing and computing devices. GPGPU's have come up in a great way in sharing the CPU's tasks and make parallel processing possible.**

***Keywords*- Spam, Bayesian Spam Filtering, Serial Spam Filter, Parallel Spam Filter, Spamicity.**

1. INTRODUCTION

Increase in internet communication has eventually led to an enormous increase of spam. Spam is unwanted data sent to a user without their wish, i.e., this data was neither asked by them nor did they desire to receive it [1]. Increase in spam leads to an enormous number of problems like slower access to emails, increase in network traffic, unwanted space occupancy and many more[2,3].

To get rid of spam two spam filters are proposed:

1. Serial Bayesian spam filter
2. Parallel Bayesian spam filter using GPGPU(general purpose computation on GPU)

The serial spam filter is designed first and later parallelized using mail list division approach to make it a parallel spam filter. In designing of both the spam filters Bayesian approach is used [4, 5, 6, 7]. The filters proposed in this paper consist of two phases: Training Phase and the Filtering Phase. In the training phase three databases are created:

- Keyword Database: Keywords are taken from the ham and spam mails
- Ham Database: Database for ham mails
- Spam Database: Database for spam mails.

The databases are taken from Enron and Snort Dataset. Once the databases are created, the spam probability of every keyword is calculated using Bayesian statistics.

The rest of this paper is organized as follows: Section 2 describes Bayesian Spam Filtering method. In this section the training phase and the filtering phase are briefly discussed. Section 3 encompasses the basics of a serial spam filter along with its design and algorithm. Sections 4 contain the basic parallel spam filter model. In this section the complete parallelization process using GPGPU is discussed. Section 5 provides a summary of serial and parallel spam filters with a summary of Bayesian spam filtering steps. We discuss the related work in Section 6 and conclude the paper in Section 7.

2. BACKGROUND

With many new filtering techniques coming up and lots of work going on in this field, daily new techniques and ways are devised to fight spam.

In work done by authors Hu Yin and Zhang Chaoyang, [8] a Bayesian serial spam filtering algorithm is implemented where the email content is tokenized and these tokens are searched for in the mail. The brute force method of searching is applied here. It is time consuming to sequentially search for each word and preprocess it separately.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Authors Amit Saxena and Mahak Motwani implemented a Bayesian spam filter in which the PFAC approach was used [9]. Here the Aho-corasick string matching is applied to calculate the count of each keyword in the email. The drawback of using this approach is that it does a dictionary match of the keywords hence deviating from the word being spam or ham. The time complexity of such an algorithm is linear and is equal to the sum total of pattern length, searched text length and no of output matches.

In work done by Monther Aldwairi and Yahya Flaifel, they have used string matching with a combination of list based and Bayesian classification to implement spam filtering [10]. This is a serial implementation. Here string matching which is approximated is till one bit level only. In case of serial spam filtering the time taken for processing mails is huge. Moreover the approximation criterion for string matching should be large to accommodate longer words.

In the system proposed for spam filtering, Bayesian statistics is applied. To search for a keyword in the email Shift-Or approximate string matching is used. This is better as compared to the exact string matching criterion because in emails spammers may change one or two letters of the word to bypass the filters. An approximate string matching criterion will have a wider look at the spam words and is able to encompass a larger arena of spam mails.

Another important and useful factor is the threshold selected or set to make the final comparison of an email and classify it as spam or ham. In this paper the threshold is trained to set itself to a value that is obtained by iteratively following the threshold setting process unless a satisfactory value is obtained. This removes the constraint of hard coding the threshold value based on historic data and makes the system more realistic by using the dataset available to calculate the threshold value.

2.1 Bayesian Spam Filtering

Bayes theorem is a mathematical formula used for calculating conditional probabilities. The complete Bayesian Spam Filtering [4, 5, 7] process consists of mainly two phases: first is the training phase and second is the filtering phase. Both the phases work in unison to obtain the result.

2.1.1 Bayesian Spam Filtering: Training Phase

This is the first phase in Bayesian spam Filtering. In the training phase keywords are extracted to make the database. Following are the steps involved:

- Historical training data is collected / taken from a source.
- Bayesian probability for all suspected keywords is calculated based on historical training data.
- Calculation of spam probabilities is done. Formula used for the same is [11,12]:

$$Spam\ Probability\ of\ a\ keyword = \frac{No.of\ occurances\ in\ spam\ dataset}{\text{No.of occurences in spam data set+No.of occurences in Ham data set}}$$

- Spamicity threshold [11, 12] is set for spam detection. The threshold can be set using historical data.

2.1.2 Bayesian Spam Filtering: Filtering Phase

This is the second phase in Bayesian spam Filtering [13]. In the filtering phase results of the training phase are used to classify mails as spam or ham. Following are the steps involved:

- Bayesian spam probabilities details are taken from the training phase
- Spam filtering Bayesian threshold details are taken from the training phase
- The suspected keyword list is loaded
- Classification of the mail is done based on the keyword occurrences in it. Classification done here is based on the training data and Bayesian network classification.

3. PROPOSED SERIAL SPAM FILTER

The proposed serial spam filter consists of two modules: the trainer module and the filter module. Considering the trainer module, there are three databases which are used in Bayesian filtering algorithm:

- Spam Suspected Keywords Set
- Trainer Ham Dataset
- Trainer Spam Dataset.

In the training algorithm spam probabilities of all the keywords are calculated based on the training dataset.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The above presented modules are designed to run on single core architecture and hence called serial spam filter. As per Bayesian statistics if we see a word "Poker" appearing in 40 of 1000 ham mails and 850 of 1000 spam mails then we can calculate the spam probability of the word "Poker " as: 850/(850+40) = 95.50 %. The training phase creates a file that contains a list of keywords with their probabilities. This is further used in the filtering phase.

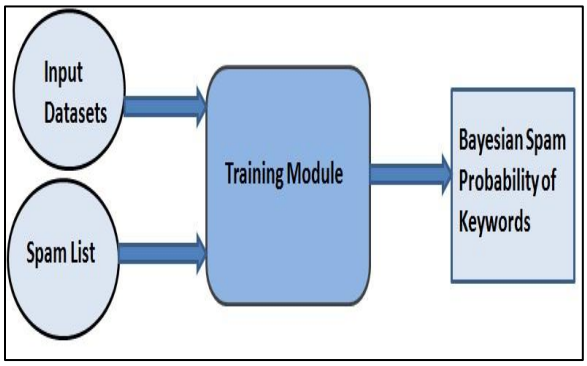


Figure 1: Training Module

Serial Spam Filter: Algorithm for Training Phase

I. Create a list of Spam Keywords and search them in ham or spam database. This list is created using data from not only one source but more than two sources.
II. Use Bayesian statistics to calculate spam probability of all keywords
III. Create a list of keywords and their corresponding probability. This file is used in the filtering module.

In the filtering module, SHIFT-OR approximate string matching algorithm is applied [14, 15]. The approximation criterion for string matching should be large to accommodate longer words. An advantage of using approximate string matching is that the spammers may change one or two bits of a word, this is done to make it bypass the filter, but approximate bit level string matching keeps a check on this.

At the input we provide list of spam keywords and test mail and get the count of all approximate spam keywords in the test mail. To determine this count we are using the SHIFT-OR algorithm. This algorithm may also count the suspected approximate keywords. Spamicity of the mail is calculated and compared with the threshold.

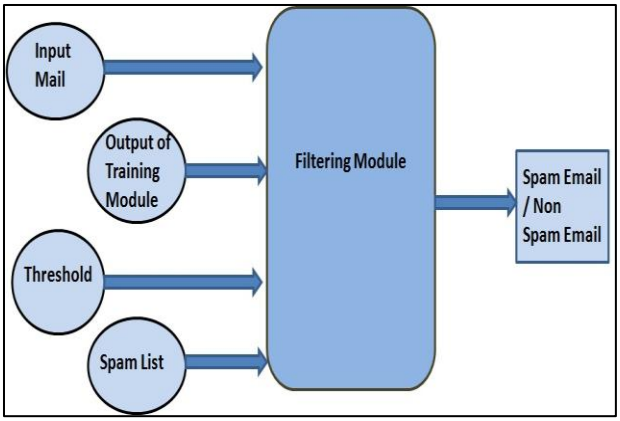


Figure 2: Filtering Module

Spamicity of the mail is calculated by using the occurrences of spam suspected keywords and the Bayesian formula given below:

$$Spamicity = \frac{p1*p2*p3\ldots\ldots Pm}{[(p1*p2*\ldots pm)+((1-p1)*(1-p2)*\ldots(1-pm))]}$$

Based on the value of spamicity obtained on comparison with the classification threshold we categorize the mail as spam or ham. This classification threshold is calculated by proposed dataset threshold training algorithm. Classification of the mail is done as per the below:

- If Spamicity >= classification threshold , the mail is reported as SPAM
- If Spamicity < classification threshold, the mail is reported as HAM

Serial Spam Filter: Algorithm for Filtering Phase

I. As one of the inputs we take list of keywords and their corresponding probabilities. This is obtained from the training phase.
II. At another input we have classification threshold.
III. Input mail is taken which is to be categorizes as ham or spam.
IV. Apply SHIFT-OR algorithm to get total frequency count of spam keywords
V. Calculate spamicity by applying Bayesian classification.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

VI. Spamicity of mail is compared with the classification threshold.
VII. If spamicity>= classification threshold, then mail is spam, else its ham.

3.1 Proposed Training Based Threshold Calculation Method

For classification of a mail as ham or spam, we have to compare the value of spamicity with a threshold value. To determine the threshold value we carry out a continuous calculation process with comparison at each iteration.

Here we are trying to calculate the classification threshold on correct values. This threshold is calculated based on the training dataset. This is an iterative process where threshold values are continuously updated for recognizing all spam mails of the dataset correctly.

We first set a threshold value based on historical data and then pass the mails as input, if the mail is identified correctly as spam or ham for the entire experimental mails database taken, then the threshold set is correct and can be used further. But if the mail classification we get is incorrect as spam and ham then the threshold value is to be changed and the whole process is to be repeated. We continue this until we get a threshold value for which we get the correct categorization of mails.

This training based threshold calculation is a unique feature of the spam filters designed. It makes the filters independent of the traditional method of using hard coded value of threshold taken from historical data. Also it makes the filters adapt and conclude a realistic threshold value obtained by training the threshold on the data available. The system is now capable of handling datasets of all kinds, since the threshold is not hard coded. It is a process that involves training of the threshold to adapt to any kind of datasets that are given at the input side of the filter.

Figure 3 shows how various mails are given to the input of the spam filter. Mail spam detector module takes these mails as input and compares the mails with its original identity. Here the decision is made whether the mail is correctly classified or not. If yes, the threshold is correct and no changes are to be made. But if the mail is not correctly identified, the whole process is repeated with a different threshold value. This process continues till a threshold value is obtained that encompassed all the items of the dataset correctly.

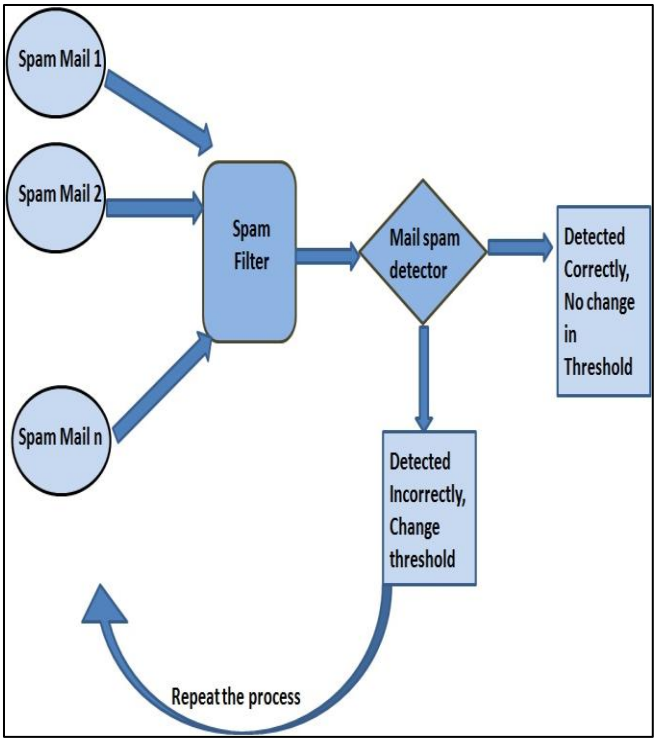


Figure 3: Threshold Setting Process

4. PROPOSED PARALLEL SPAM FILTER ON GPGPU'S

GPGPU's are great for data parallelism [16]. They are designed for tasks that are highly parallel. The GPU is ALU heavy [17], i.e., there is lots of compute power available. The applications targeted to be run on GPU should have:

- High parallelism
- High arithmetic intensity
- Minimum dependency between data elements
- Huge data sets
- Lot of work to be done without CPU intervention

The serial spam filter discussed in the previous section is parallelized. This is done by using mail division method on GPGPUs. The parallel spam filter [18] consists of two modules: Trainer module and Filter module.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

4.1 Parallel Spam Filter: Trainer Module

In the trainer module three databases are created , as in case serial spam filter: spam keywords database, ham database and spam database. These databases are then made to pass the Bayesian Training Algorithm that calculates the spam probability of all spam keywords depending on the training dataset taken. The training module here is not parallelized because the whole process of training takes place only once.

Parallel Spam Filter: Algorithm for Training Phase

I. Take Spam Keywords from the standard data libraries like Enron, Snort and Wikipedia and form a list.
II. Search the listed keywords in ham or spam database.
III. Calculate the spam probability of each keyword through Bayesian statistics.
IV. Store the keywords and their spam probabilities in a separate file.
V. This file is used in the filtering module.

4.2 Parallel Spam Filter: Filtering Module

In the filtering module a parallel algorithm is devised which takes spam keywords and tests multiple mails on multiple cores as input. There is a core-mail coupling, i.e., cores are assigned to mails. Each core is given a test mail and the count of keywords in that mail is obtained. After this spamicity of all the mails are calculated simultaneously by using the Bayesian Formula. At the outputs we have results of multiple mails as spam or ham. This is done by comparing the spamicity with the threshold, if spamicity>= threshold its spam else ham. All the work is done on multiple test mails in parallel.

There could be various ways of passing the mails in the proposed system to make its parallel version. Since data is involved and can be toggled in many ways, these possibilities arise. We have applied the concept of mail division wherein all the mails of the dataset are given to n cores simultaneously and all the n cores run a copy of the spam filter to filter and identify the mail as spam or ham. It provides considerably large efficient time utilization.

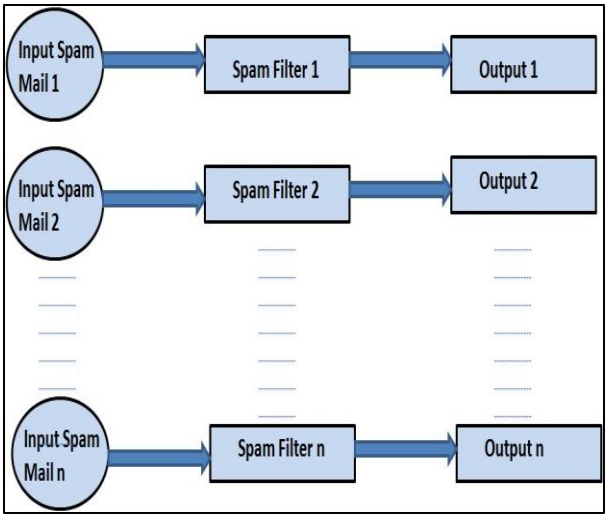


Figure 4: Parallel Spam filter

Parallel Spam Filter: Algorithm for Filtering Phase

I. Fetch the file created in the training module containing the keywords and their probabilities.
II. Fetch the test mails.
III. Mail division method is applied to mails for calculation of frequency count of spam keywords in each test mail.
IV. Comparison of spamicity with threshold is done.
V. Declaration of mail as spam or ham.

5. EXPERIMENTAL SETUP

Number of spam suspected keywords in our case are 196 taken from Wikipedia and Snort Dataset. Data sets used in the training are Enron and snort..In the Training module of serial Bayesian Filter, count of each keyword in ham and spam is calculated and stored in a .txt file called spam-probability file. The filter module calculates the spamicity of an input mail by calculating the count of each keyword appearing in the mail.

All the experiments were performed on Intel(R) Core(TM) i7 CPU @2.4GHz. Other resources include primary memory (RAM) 8 GB, 64-bit version of Windows7 operating system, visual C++ runs on visual studios 2008 to version 2012. 70% time of the CPU is used as execution time for heavy loaded system. . Each experiment is carried out 10 times in order to take average of all times. For all the experiments, 1000 different emails approximately of size 0.5 MB each are taken.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Serial Spam Filter:
After the spam probability file is created and populated, below screenshot is obtained:

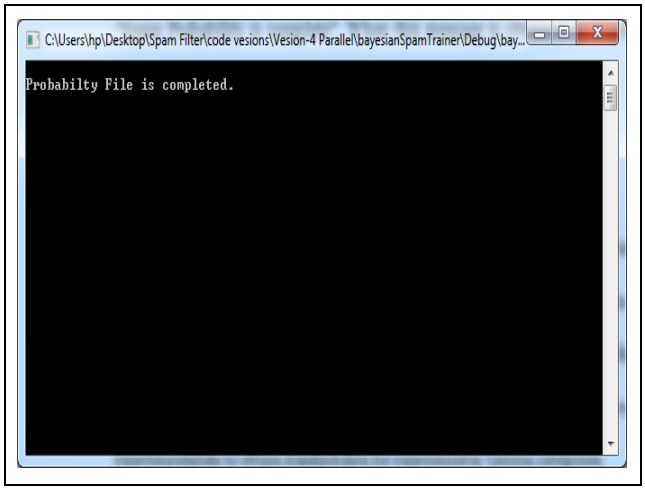


Figure 5: Spam Probability File Creation

On detection of Ham or Spam mails we get the below screenshots:

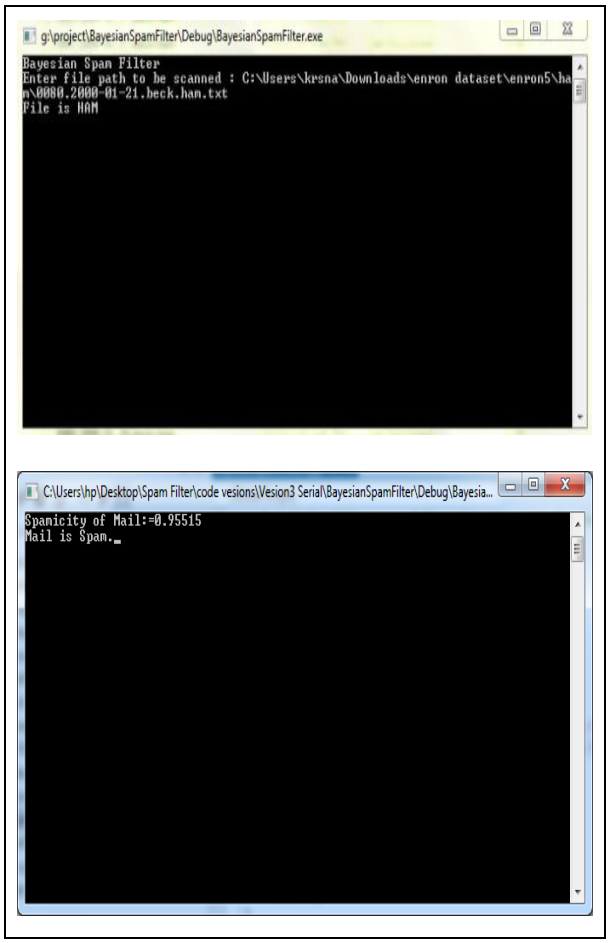


Figure 6: Declaration of Mail as Ham or Spam

Result: Successful detection of spam mails with accuracy of 93%. We tested 2000 mails out of those approximately 1857 mails were successfully detected. These mails were a database of spam mails and the same for tested specifically for spam mails.

Parallel Spam Filter:
This parallel spam filter is implemented on GPGPU's using the OpenCL programming language [19,20].

Processor: Core i3, RAM: 4 GB, OS: Windows 7, Language: Visual C++ runs on Visual Studios 2008, GPGPU: AMD Radeon HD 6800 series(192 cores and 960 processing threads), Language (parallel implementation): OpenCL.

6. EXPERIMENTAL RESULTS & ANALYSIS

The results are calculated and compared based on the following criterion: Speedup, Accuracy, Recall, Precision and F- Measure.

The speedup is in terms of the time taken by the spam filters to execute a chunk of mails. Supporting graphs and tables are presented in the ongoing section.

Below are the speedup table and graph comparing the speedup of serial and parallel spam filter implemented in the previous sections.

TABLE 1: SPEEDUP TABLE OF PARALLEL FILTER SPEED AND SERIAL FILTER SPEED(IN SECS)

S.No.	No. of Test Mail	Parallel Filter Speed	Serial Filter Speed
1.	1000	0.18	130
2.	2000	0.23	190
3.	5000	0.48	310
4.	10000	1.14	609

Various chunks of test mails are passed to the filter: 1000 mails, 2000 mails, 5000 mails and 10000 mails. The speedup is then calculated by tracking the time taken for the complete execution to take place. Comparing the results of the serial and parallel filters implemented we see that there is remarkable improvement in the time taken by the serial and parallel filters as expected.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

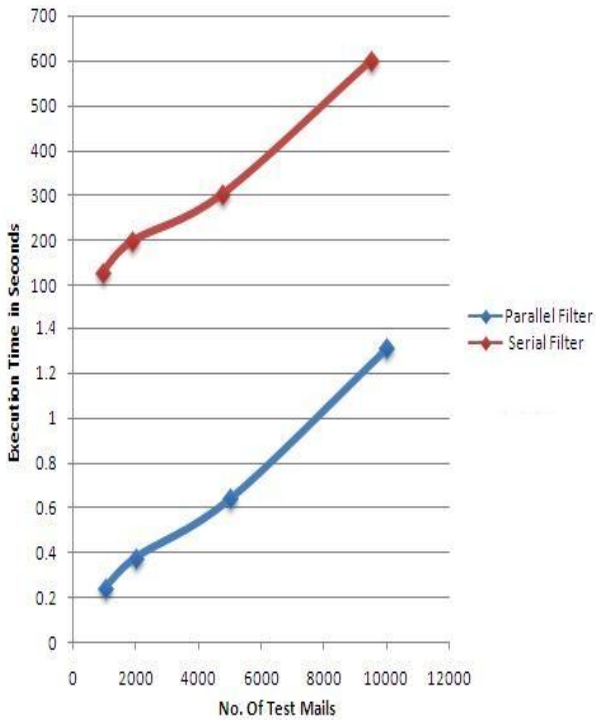


Figure 7: Comparison Graph of Serial and Parallel Spam Filter

The above graph depicts the comparison of speedups between the serial and parallel spam filter on the x-axis no of test mails is taken and on y-axis time is taken in seconds. This is a bi-scalar graph that is plotted due to the huge variation between the timings of serial and parallel filters.

TABLE 2: ACCURACY MEASUREMENT TABLE

S.No.	No. of Test Mails	Accuracy
1.	1000	93%
2.	2000	91 %
3.	5000	92 %
4.	10000	90 %

From the above table it's evident that the spam filter's accuracy rate is approximately 90%. For training we have taken Enron and Snort data sets. Keywords are limited in our experiments and can be increased. If we will increase training data sets and keywords than results will be more accurate and efficient.

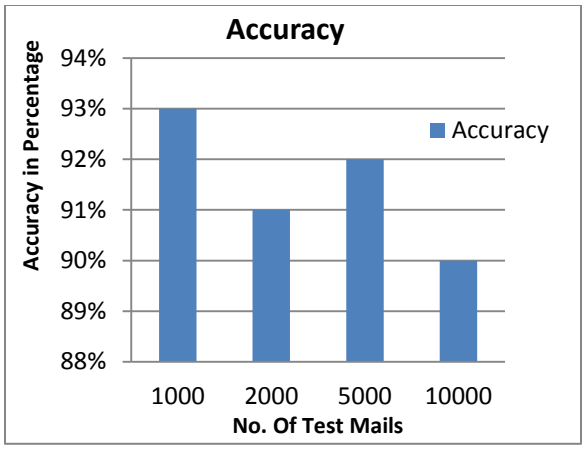


Figure 8: Accuracy of Proposed Parallel Spam Filter

In Figure 8 above, no of test mails is taken on x-axis and accuracy in percentage is taken on the y-axis. Chunks of mails are analyzed and the accuracy is plotted accordingly. Considering it on an average we get the accuracy of the system somewhere close to 90% which is better than the accuracy achieved on the experiments conducted earlier.

Taking it to the next level of comparisons, Speedup is compared with the existing parallel version of the Bayesian spam filter where a different technique was applied to parallelization.

Speedup is now compared for the existing PFAC based parallel spam filter and our proposed spam filter based on mail division method. Speedup is compared as per the test results obtained from the training data and experimental results. We see a marginal increase in speedups.

TABLE 3: SPEEDUP TABLE OF PROPOSED PARALLEL FILTER SPEED AND EXISTING FILTERS SPEED (IN SECS)

S.No.	No. of Test Mails	Proposed Parallel Filter	Existing Parallel Filter Speed
1.	1000	0.18	0.244
2.	2000	0.23	0.38
3.	5000	0.48	0.64
4.	10000	1.14	1.31

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

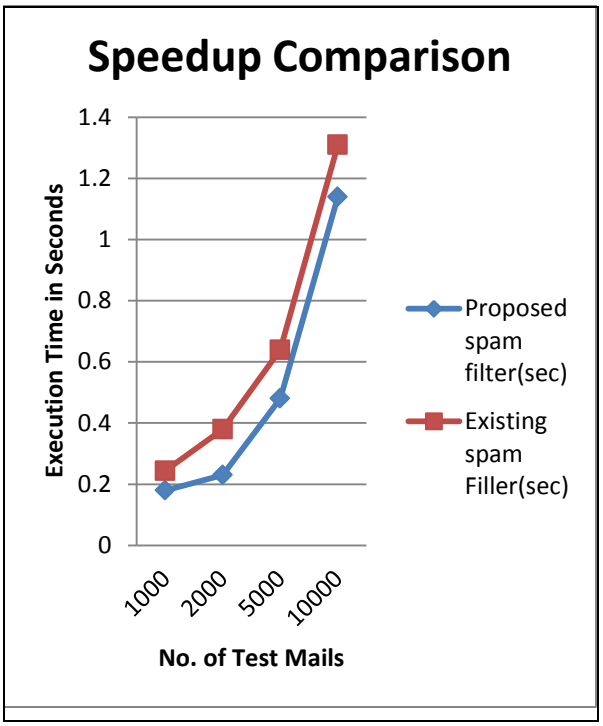


Figure 9: Comparison Graph of Proposed and Existing Spam Filter

Figure 9 shows the speedup comparison between the existing PFAC based spam filter and proposed mail division method based spam filter. There is a marginal increase in the speeds when tested on the chunks of mails and passing it to 1000 cores.

The accuracy of the existing PFAC based parallel spam filter [9] is compared with the accuracy we received by our mail division method based parallel spam filter. The accuracy has remarkably improved and we were able to get better results that were more exhausted and give better performance.

TABLE 4: ACCURACY MEASUREMENT TABLE

S.No	No. Of Test Mails	Our Experimental Accuracy	Existing Setups Accuracy
1.	1000	93%	76%
2.	2000	91%	72%
3.	5000	92%	70.9%
4.	10000	90%	70.1%

We are able to achieve a considerable improvement in the accuracy of our proposed parallel spam filter. There is an improvement of 15% - 20% in the accuracy of the system. Where the average accuracy of the present PFAC based parallel spam filter was 70%, we achieved an accuracy of 90% in the results obtained using mail division method

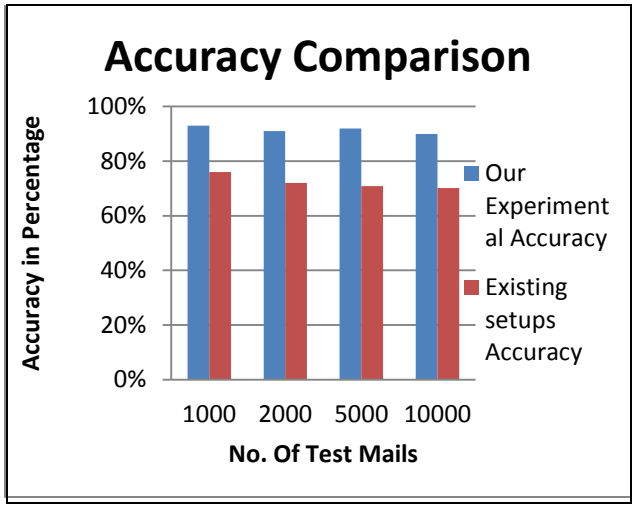


Figure 10: Accuracy Comparison of Proposed and Existing Parallel Spam Filter

In figure 10, the existing setup is depicted in pink and our proposed setup in blue. As it's evident from the graph, there is improvement in accuracy of the system proposed.

Another set of evaluation parameters are precision, recall and f-measure. We do the analysis of these parameters also based on our experimental results obtained [21, 22].

For the calculation of precision, two datasets are considered: Snort dataset with a total of 5000 mails and Enron dataset with a total of 2000 mails. These mails are a combination of ham and spam mails. For the calculation of Recall, 20 high probability spam keywords are removed from the keyword list and then the recall is calculated on a dataset of 1000 mails each from Snort and Enron. The recall is obtained without the use of these keywords, hence it becomes tougher for the filter to identify the mail as spam or ham, but we achieved considerably good results which show the accuracy and efficiency of our spam filter.

TABLE 5: PRECISION, RECALL AND F- MEASURE EVALUATION TABLE

	Snort Dataset	Enron Dataset
Precision	94.67 %	98.7 %
Recall	92.37 %	97.13 %
F-Measure	93.50%	97.90%

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Considering the above results we see that high precision and recall is obtained for the dataset where mails were taken from Enron database. The values of precision as 98.7 % is a very good result obtained that interprets to a very effective performance by the filter. Similarly the value of recall which is generally considered to be not obtained that well is also very good in our case 97.13 %.

7. CONCLUSION & FUTURE WORK

The proposed spam filters are tested for a variety of different mails which were categorized as spam or ham. Data was taken from Enron and Snort data set. Results are explained with the help of graphs and the results shows that as the no. of mails will increase, efficiency will decrease and after certain point efficiency will remain almost same. Efficiency of our spam filters was approximately 90%. The parallel spam filter is significantly faster than serial spam filters. This speedup is directly based on SIMD parallel architecture and number of mails supplied in the unit time.

In future work, one can continue further research and refinement. The refinement can be on keyword selection, to further enhance overall performance of spam Bayesian filtering. More accurate information about the keywords can provide better threshold calculation and better spamicity calculation. Also there could be an improvement in the efficiency of spam filter with different SIMD architecture. The accuracy of the spam filter can be further enhanced by improving the Shift-OR approximation criteria or using another variant like Q-gram.

8. REFERENCES

[1] Yanti Rosmunie Bujang , Husnayati Hussin "Should We Be Concerned with Spam Emails?
A Look at Its Impacts and Implications", presented at the *5th International Conference on Information and Communication Technology, IEEE 2013,* p 01.

[2] Izabella Miszalska, Wojciech Zabierowski, Andrzej Napieralski, "Selected Methods of Spam Filtering in Email, ", CADSM'2007, February 20-24, 2007, Polyana, UKRAINE ,p 02
\

[3] "Spam email percentage in mailbox" McAfee Managed Mail Protection Always On, Automatic Mail Protection. Available: Website:http://www.mcafee.com/us/resources/misc/web-protection-infographic.pdf [Accessed: Feb 5, 2014]

[4] Tiago A. Almeida and Akebo Yamakami, "Content-Based Spam Filtering", IEEE, 2010

[5] T. Guzella and W. Caminhas, "A review of machine learning approaches to spam filtering," Expert Systems with Applications, 2009, in press.

[6] Y. Song, A. Kolcz, and C. L. Giles.,"Better naive bayes classification for high precision spam detection", Softw. Pract. Exper., 39:1003–1024, August 2009.

[7] Phimphaka Taninpong, Sudsanguan Ngamsuriyaroj "Incremental Adaptive Spam Mail Filtering Using Naïve Bayesian Classification", 2009 10th ACIS International Conference on Software Engineering, Artificial Intelligences, Networking and Parallel/Distributed Computing.

[8] Hu Yin, Zhang Chaoyang, "An improved Bayesian Algorithm for Filtering Spam E-mail", IEEE, 2011 International Symposium on Intelligence Information Processing and Trusted Computing p 02

[9] Amit Saxena, Mahak Motwani, Saima Haseeb, "Serial and Parallel Bayesian Spam Filtering using Aho-Corasick and PFAC", Volume 74– No.17, July 2013, International Journal of Computer Applications, pp-5,6.

[10] Monther Aldwairi nad Yahya Flaifel, "Baeza Yates and Navarro Approximate String Matching for Spam Filtering ", IEEE, 2012, p 18

[11] C. Pu, S. Webb, O. Kolesnikov, W. Lee, and R. Lipton. Towards the Integration of Diverse Spam Filtering Techniques. In Proc. of IEEE International Conference on Granular Computing, pages 7 – 10, 2006.

[12] I. Androutsopoulos and et., "An Evaluation of Naïve Bayesian Anti-Spam Filtering", 11th EurpoeanConference on Machine Learning, pp 9-17, Barcelona, Spain, June 2000.

[13] Robert Haskins and Rob Kolstad"Bayesian Spam-Filtering Techniques",The Advanced Computing Systems Association & The System Administrators Guild, volume 28,number 3,June 2003.

[14] Gonzalo Navarro and Mathieu Raffinot. "A Bit Parallel approach to Suffix Automata: Fast Extended String Matching",In M. Farach (editor), Proc. CPM'98, LNCS 1448. Pages 14-33, 1998.

[15] Rajesh Prasad, Suneeta Agarwal, Ishadutta Yadav, Bharat Singh, "Efficient Bit-Parallel Multi-Patterns String Matching Algorithms for Limited Expression",ACM Jan 22-23, 2010.

[16] Jingweijia Tan , Yang Yi , Fangyang Shen , Xin Fu
"Modeling and characterizing GPGPU reliability in the presence of soft errors" Parallel Computing 2013, ELSEVIER IEEE 2007, pp 01-02, 521-523.

[17] He Bingsheng, Huynh Phung Huynh,R.G.S. Mong,"GPGPU for real-time data analytics", In the proc. IEEE 18th International Conference onParallel and Distributed Systems (ICPADS), pp. 945-946, Dec. 2012.

[18] Xiao-Ping Liu, En-zhu Wang, Li-ping Zheng, Xing-wuWei,"Study on Template for Parallel Computing in Visual Parallel Programming Platform" in the proc. of 1st International Symposium on Pervasive Computing and Applications, pp. 476-481, Aug. 2006.

[19] JieShen, Jianbin Fang, H. Sips, A.L. Varbanescu,"Performance Gaps between OpenMP and OpenCL for Multi-core CPUs", In the proc. of 41st International Conference onParallel Processing Workshops (ICPPW), pp. 116-125, Sept. 2012.

[20] The Khronos OpenCL Working Group, "OpenCL - The open standard for parallel programming of heterogeneous systems."

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Weblink: http://www.khronos.org/opencl [Accessed on April 5, 2014].

[21] Kunal Mehrotra Shailendra Watave, Spam Detection, A Bayesian approach to filtering spam.pp 9-15

[22] M. Tariq Banday, Tariq R. Jan "Effectiveness and Limitations of Statistical Spam Filters", International Conference on "New Trends in Statistics and Optimization" Department of Statistics, University of Kashmir, Srinagar, India, from 20th to 23rd October,2008, pp-9-13

Mobile - Health Application Software Design and Development

Ayangbekun Oluwafemi J.
Department of Information Systems
University of Cape Town
South Africa

Kasali Olanrewaju M.
Department of Information Technology
Crescent University Abeokuta, Nigeria

***Abstract*— Mobile technologies are fast developing and it has completely changed the way we interact and provide healthcare services. The rapid spread of mobile technologies and inventive applications to address health related problems has evolved into a new field known as mobile-Health. The purpose of this research is to improve the quality and access to health care services with the aid of mobile-Health application software known as "Crescent Mobile Health". This paper will address the problem of self medication by creating a channel of communication between a patient and doctor at distant environment there by solving emergency situations. The method used to address this problem is by designing and developing mobile-Health application software, which can be used by patients via an android smartphone that is used to communicate with a doctor/pharmacist/laboratory scientist using electronic-Health application software known as Crescent Health Information System on a desktop via the intranet. The two applications on smartphone and desktop are able to communicate via instant messaging by a persistent connection known as "sockets" and "pusher" which provides implementation for interconnectivity.**

The Crescent Health Information System can carry out major functionalities such as drugs and tests inventory, instant messaging, prescriptions of drugs, prescription of tests and profile update. The Crescent Mobile Health can also carry out functionalities such as instant messaging, viewing of prescribed drugs, tests, health tips and help file. The mobile-Health application software was developed using java programming language and android development studio while the electronic-Health (E-Health) application software was developed using PHP programming language and MYSQL database. The results of the development of this project concludes that mobile-Health application software has been able to resolve the problem of communication between a patient and a doctor and has provided a means to verify drugs available and tests carried out in the clinic/health sector.

Keywords - Electronic-Health; Healthcare; Intranet; Mobile-Health; Patient; Smartphone; Socket

I. INTRODUCTION

Mobile-Health (M-Health) is referred to using mobile communication devices, such as mobile phones, tablet, computers and PDA (Personal Digital Assistant) phones for health services and information [1]. The unprecedented spread of mobile technologies as well as advancements in their innovative application to address health priorities has evolved into a new field of electronic-Health, known as mobile-Health. The mobile-Health field has emerged as a sub-segment of electronic-Health; the use of information and communication technology (ICT), such as computers, mobile phones, communications satellite and patient monitors for health services and information [2].

However, Mobile-Health application software includes the use of mobile devices in collecting community and clinical health data, delivery of healthcare information to practitioners, researchers, and patients, real-time monitoring of patient vital signs, and direct provision of care (via mobile telemedicine) through application of software developmental paradigms [3].

Patients are often tired of waiting to be attended to and are often shy of disclosing some medical conditions to the doctors face to face. But with the advent of mobile-Health application software, each patient can sit at their comfort zone and get diagnosed easily in a secured and confidential environment and can also verify the availability of drugs and various medical tests facilities available in the clinic without being physically present.
Therefore, this research addresses:

- How to refine the mobile-Health service application model to make it suitable for a patient to understand.
- How to encourage patients to stop self medication.
- How to encourage communication between a patient and a doctor.
- How to modify and improve health services.
- How to encourage fast transfer of information concerning health epidemics and health tips.

II. MOBILE-HEALTH (M-HEALTH) CONCEPT

The rapid expansion of mobile information and communications (ICT) technologies within health service delivery and public health systems has created a range of new opportunities to deliver new forms of interactive health services to patients, clinicians, and caregivers alike [4]. Mobile technologies can include, but are not limited to, tablets, cell phones (hardware and software) and Smartphone,

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

mobile-enabled diagnostic and monitoring devices, or devices with mobile alert systems. Mobile-Health can be referred to as the segment of healthcare delivery broadly defined as health-related services to patients, clinicians, and caregivers through mobile technology platforms on cellular or wireless networks [5].

Early in its development, in 2003, mobile-Health was defined as wireless telemedicine involving the use of mobile telecommunications and multimedia technologies and their integration with mobile healthcare delivery systems [6]. Since then it has come to encompass any use of mobile technology to address healthcare challenges such as access, quality, affordability, matching of resources, and behavioural norms. Thus it can involve a wide variety of people and products, as well as the actions that connect them. The crux of these connections is the exchange of information. Mobile technologies cannot physically carry drugs, doctors, and equipment between locations, but they can carry and process information in many forms: coded data, text, images, audio, and video [7].

III. MOBILE TECHNOLOGIES IN HEALTHCARE

The adoption of mobile technologies in every dimension of life has been phenomenal. In the span of two decades, ever-more sophisticated mobile technology has fundamentally altered the ways in which people communicate and conduct business. The disruptive power of these new technologies and the accompanying waves of innovation they have sparked are transforming the health care industry, propelling stakeholders to reassess and repurpose how they provide services. The capabilities offered by mobile technologies are fast becoming appreciated by industry stakeholders, with a raft of devices, sensors, apps, and other programs being developed that target chronic conditions, telemedicine and remote monitoring, patient data capture, electronic records, e-prescribing, and the parallel industries of fitness and wellness [8].

IV. RELATED WORK

A. *Gazelle Mobile-Health Application by Quest Diagnostics Inc*

The mobile-health application that was developed by Quest Diagnostics Inc known as 'Gazelle', allows users to receive their Quest Diagnostics lab results and manage their personal health information directly from their Blackberry, Apple iPhone, Google Android, Smartphones. This application enables users to see, store and share their vital health information with ease and security while on the go. Patients can request and receive their Quest Diagnostics laboratory results directly to their smartphone. Gazelle allows a patient to easily share up-to-date medical information. Laboratory test results can be conveniently e-mailed or faxed directly from a smartphone to physicians or other caregivers. In the event of an emergency, first responders can gain immediate access to critical healthcare information [9].

B. *Itriage Mobile-Health Application by Aetna and Denver-Based Itriage*

Denver-based iTriage, maker of a mobile and Web app that helps consumers find physicians based on symptoms they are experiencing, is working with parent company Aetna to offer mid-sized employer groups a customizable version of the app that guides users to in-network providers. iTriage combines health information with GPS and mapping technology to help you find care whether you are travelling, at work or close to home. Description Created by two ER docs, iTriage helps you answer the questions: "What medical condition could I have?" and "Where should I go for treatment?" Save, easily access, and share the healthcare information that's most important to you [10].

C. *Webmd Mobile-Health Application by Webmd Health Corp*

WebMD for Android helps with decision-making and health improvement efforts by providing mobile access 24/7 to mobile-optimized health information and decision-support tools including WebMD's Symptom Checker, Drugs & Treatments, First Aid Information and Local Health Listings. WebMD the App also gives you access to first aid information without having to be connected wirelessly – critical if you don't have Internet access in the time of need [11].

V. METHODOLOGY

The proposed system will be implemented using Java programming language and android development studio which is suitable for the development of mobile-Health application software. The mobile-Health application software will be integrated with electronic-health application software, which will allow for communication to take place and access to the database will be granted through this means. The electronic-Health application software which is a web application will be implemented using PHP and MYSQL database.

Unified Modelling Use Case diagram will be used to depict the basic functions of the proposed system. Waterfall Model will be employed as the Software development process model. The model will be employed in order to have an efficient and effective workflow for the development of the mobile-Health application software.

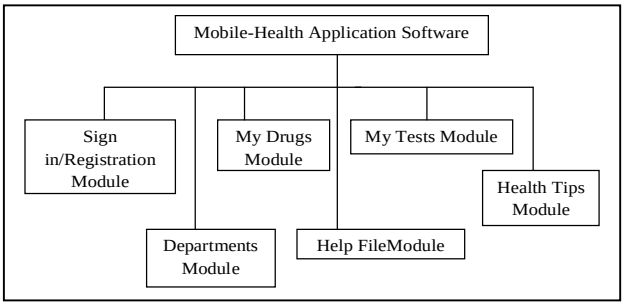


Figure 1. Hierarchy of mobile-health application software

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The following are the mobile-Health application software sub-modules (Fig. 1) and their descriptions;

- **Sign in/Registration Module** is where new patients can register and old patients can sign into the application. Information such as full-name, email address, gender, blood group, address and password will be provided here (Fig. 11).
- **Departments Module** contains information about various departments in the hospital and the doctors/pharmacists/laboratory scientists available in each of the departments.
- **My Drugs Module** is where the drugs that have been prescribed to the patient by the doctor can be viewed (Fig. 13).
- **Help File Module** contains information about the definition of various modules in the mobile-Health application software and also solutions to the challenges a user can encounter while using the application.
- **My Tests Module** contains information about test that has been prescribed to the patient by the doctor.
- **Health Tips Module** contains healthcare information that is of advantage to the patient.

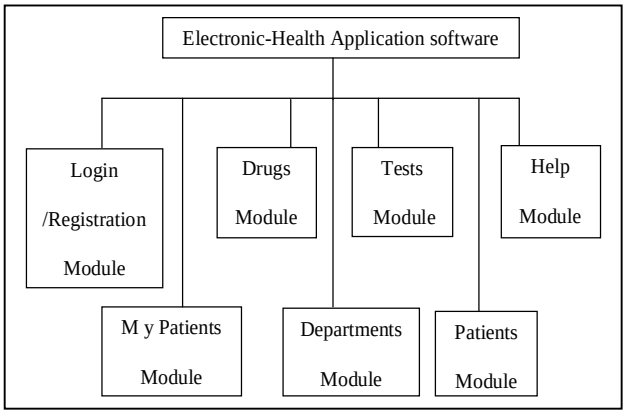


Figure 2. Hierarchy of electronic-health application software

However in term of an electronic-Health application software developmental approach/hierarchy (Fig. 2), the following are the proposed systems sub-modules and their descriptions:

- **Login/Registration Module** allows authorized doctors/pharmacist/laboratory scientist to log into the web application to in order to gain access to the application and also new doctors/pharmacist/laboratory scientist are able to register into the application (Fig. 8).
- **My Patients Module** displays the patients that are available for the chat session to be initiated. From here, the doctor can select a patient to chat with, prescribe drugs and tests for the patient and also make case notes concerning the patient's health. The pharmacist and laboratory scientist can chat with patients through this module to verify drugs and tests (Fig. 9).
- **Drugs Module** is where drugs available in the clinic can be viewed. The drugs module contains information such as the name of the drug, price, category, status, description and editing of the drugs.
- **Departments Module** contains information about the departments available in the hospital. The present department can be edited and also new departments can be added to the application.
- **Test Module** displays information about the test that can be carried out in the hospital laboratory. Test modules contains information such as name of the test, price, status, description, department and editing of the tests.
- **Patients Module** contains information about patients that has registered into the mobile-Health application software. Information such as username, status, name, blood group, gender, date registered and actions are found here. Dormant patients can also be suspended (Fig. 10).
- **Help Module** contains information about challenges a user can encounter while using the application and the solution to the challenges.

VI. DESCRIPTION OF THE PROPOSED SYSTEM

For the mobile-Health application software to be functional, it has to be integrated with electronic-Health application software on a desktop. The electronic-Health application software will be installed on the doctor's desktop which will be used to communicate with the patient on the mobile-Health application software on an android smartphone.

A. M-Health Application Software on a Smartphone

The mobile-Health application software's first interface requires the user (patient) to provide email address and password to access the application. If the user has not registered into the application, the user will click on 'Go to register' instead of 'Sign in' (Fig. 11), after clicking on go to register the user will have to provide information such as full name, email address, blood group (where the user will be given a list of blood group to choose from), gender, nationality, address, password and confirm password. Unregistered users can view health tips and help file by pressing the option button on the android smartphone.

The second interface after the user has signed in is the departments interface. From the departments interface, the user can choose the department of one's choice and communicate by chatting with the medical personnel in that department (Fig. 12). User can then view the drugs (Fig. 13) and test that has been prescribed by the doctor. The user can also view health tips which are of benefit to the user's health and help file which will assist the user in any challenge the user may encounter while using the application. The mobile-Health application software (Crescent Mobile Health) smartphone is an Android application.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

However, all updates concerning departments, doctors, pharmacists and laboratory scientists will be automatically updated on the mobile-Health application software. Meanwhile, all updates concerning patients will be automatically updated on the electronic-Health application software.

B. E-Health Application Software on a Desktop

In the electronic-Health application software, the first interface requires the user (doctor, pharmacist or laboratory scientist) to provide access name and password to access the application. If the user has not registered into the application, the user will click on 'Register' instead of 'Login' (Fig. 8), in which after clicking on register the user will have to provide information such as user name, full name, department (where a list of departments is displayed for the user to choose from) and password.

After the user has registered into the application and has logged into the application, the user will be taken to the second interface which will consist of six main (Fig. 2) modules such as my patients, drugs, department, test, patients (Fig. 10) and help. My patients module which is the second interface the user will encounter will provide a channel where the doctor can chat with the patients, make case notes, prescribe drugs and tests (Fig. 9).

VII. UML USE CASE DIAGRAM

The Unified Modeling Language (UML) is a data modeling technique that depicts the actors of a system and the actions they can perform in the system via a diagram called the Use Case diagram. Below are The Use case diagrams (Fig. 3 and Fig. 4) depicting the actors "Doctors", "laboratory scientists", "pharmacists" and "Patients" of mobile-Health application software and electronic-Health application software and the respective actions they can perform.

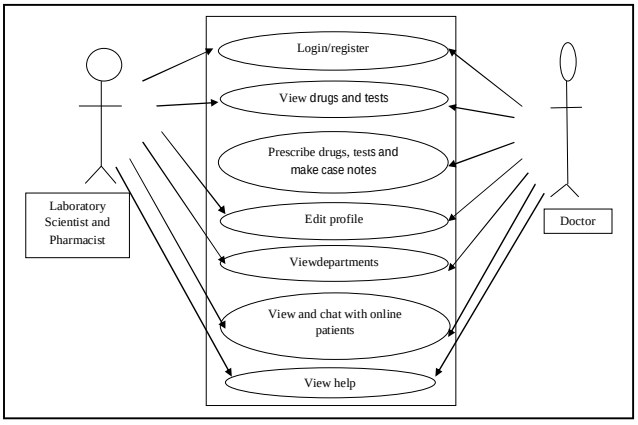


Figure 3. Use case diagram for doctor, laboratory scientist and pharmacist using electronic-health application software

The Use Case diagram (Fig. 3) above describes the proposed roles attached to a doctor/laboratory scientist/pharmacist that will use the electronic-Health application software on a desktop.

The roles are:

- **Login/register:** The doctor, pharmacist and laboratory scientist all have the ability to either register as a new user or login as an existing user of the application.
- **View drugs and tests:** Doctor, pharmacist and laboratory scientist can all view the drugs and tests available in the clinic.
- **Prescribe drugs, tests and make case notes:** Only doctors can prescribe drugs, tests and make case notes through my patients' module of the electronic-Health application software.
- **Edit profile:** Doctor, laboratory scientist and pharmacist can edit their profile on the electronic-Health application software.
- **View departments:** Doctor, laboratory scientist and pharmacist can view patients that registered into the mobile-Health application software.
- **View and chat with online patients:** Doctor, laboratory scientist and pharmacist can view online patients and chat with them.
- **View help:** Help file can be viewed by doctor, laboratory scientist and pharmacy to help provide solutions to challenges they might encounter.

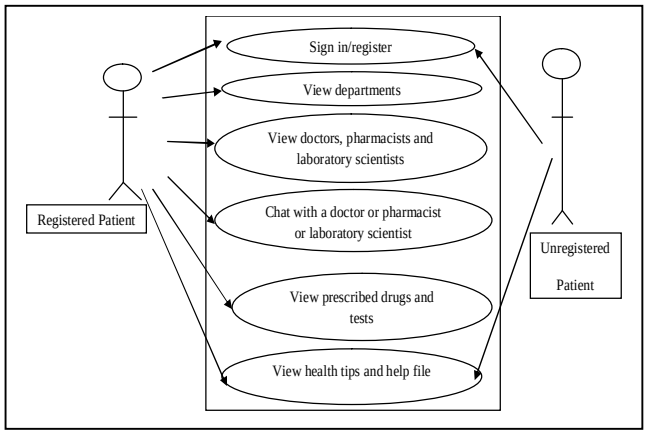


Figure 4. Use case diagram for a registered patient and unregistered patient using mobile-health application software

The Use Case diagram (Fig. 4) above describes the proposed roles attached to a patient that will use the mobile-Health application software on a smartphone.
The roles are:

- **Sign in/register:** A new patient must register to use the application. But an old patient will only sign in to access the functionalities of the application. Unregistered patient can only view health tips and help file by pressing option button on the android smartphone.
- **View departments:** A patient who has signed in can view all the departments available on the electronic-Health application software.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

- **View doctors, pharmacists and laboratory scientists:** A patient can view doctors, pharmacists and laboratory scientist of his or her choice pertaining to his or her area of interest.
- **Chat with a doctor or pharmacist or laboratory scientist:** A patient can chat with a doctor or laboratory scientist or pharmacist of one's choice.
- **View prescribed drugs and tests:** If a drug or test has been prescribed to a patient by a doctor via electronic-Health application software, the patient can view the drug or test that has been prescribed to him or her via mobile-Health application software.
- **View health tips and help file:** A patient can view health tips and help file on the mobile-health application software.

VIII. ARCHITECTURE OF AN ANDROID OPERATING SYSTEM

Android is an open source and Linux-based "Operating System" for mobile devices such as smartphones and tablet computers. Android operating system is a stack of software components which is roughly divided into five sections and four main layers as shown below in the architecture diagram (Fig. 5).

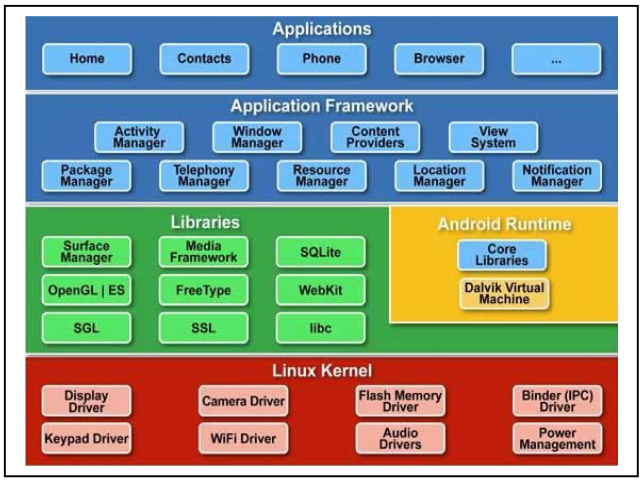


Figure 5. Architecture of an android operating system [12]

- **Applications**
  The android application is located at the top layer. The mobile-Health application software will be written to be installed on this layer only.
- **Application Framework**
  The Application Framework layer provides many higher-level services to applications in the form of Java classes. Application developers make use of these services in their applications.
- **Libraries**
  On top of Linux kernel there is a set of libraries including open-source Web browser engine WebKit, well known library libc, SQLite database which is a useful repository for storage and sharing of application data, libraries to play and record audio and video, SSL libraries responsible for Internet security etc.
- **Android Runtime**
  This is the third section of the architecture and available on the second layer from the bottom. This section provides a key component called **Dalvik Virtual Machine** which is a kind of Java Virtual Machine specially designed and optimized for Android. The Dalvik VM makes use of Linux core features like memory management and multi-threading, which is intrinsic in the Java language. The Dalvik VM enables every Android application to run in its own process, with its own instance of the Dalvik virtual machine. The Android runtime also provides a set of core libraries which enable Android application developers to write Android applications using standard Java programming language.
- **Linux kernel**
  At the bottom of the layers is Linux - Linux 2.6 with approximately 115 patches. This provides basic system functionality like process management, memory management, device management like camera, keypad, display etc. Also, the kernel handles all the things that Linux is really good at such as networking and a vast array of device drivers, which take the pain out of interfacing to peripheral hardware [12].

IX. INTEGRATION OF ELECTRONIC-HEALTH AND MOBILE-HEALTH APPLICATION SOFTWARE

The mobile-Health application software is connected to the database via Application Programming Interface (API) and this connection is passive (the transfer of information is not simultaneous). API specifies how some software components should interact with each other. In addition to accessing databases or computer hardware, such as hard disk drives or video cards, an API can be used to ease the work of programming graphical user interface components [13].

For the electronic-Health application software and mobile-Health application software users to be able to communicate (peer-to-peer connection) via instant messaging, a persistent connection is required. Sockets, which is persistent/ active (concurrent transfer of information) for peer-to-peer communication is used. Pusher which is a hosted API for quickly, easily and securely adding scalable real-time functionality to web and mobile applications acts as a third party application that provides an implementation of sockets, thus facilitating peer-to-peer communication. Thus immediate transfer and receiving of information occurs via this integration.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

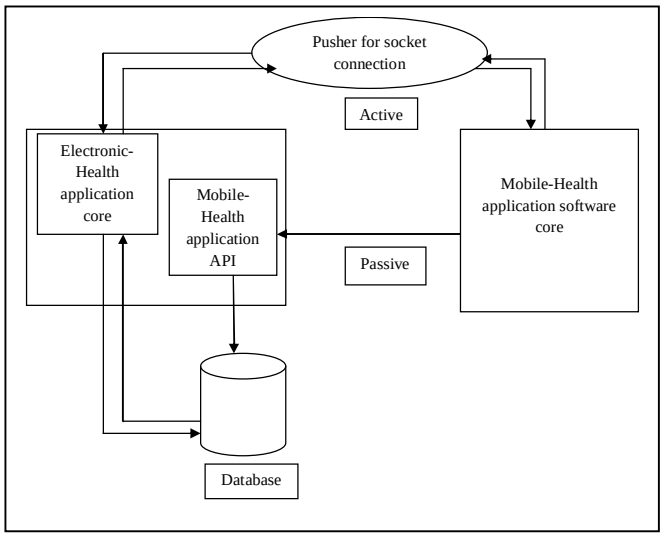


Figure 6. Integration of e-health application software and m-health application software diagram

The diagram (Fig. 6) above describes how electronic-Health application core fetches from and stores data in the database, how mobile-Health API fetches from and stores data in the database. It also explains how mobile-Health application software core links with mobile-Health application API to communicate with the database via predefined API functions.

Note: Mobile-Health application software and electronic-Health application software make use of the same database thereby facilitating rapid updates between mobile-Health application software and electronic-Health application software.

X. DATABASE DESIGN

The database system employed for this application is MySQL Relational Database server (a Relational Database Management System). This Database server stores and retrieves data by the presentation layer. A Relational Database Management System was selected as the choice for the database design because of its advantages which includes: consistency of records, data duplication avoidance, easy modification and manipulation of data and data format and better security. The proposed system (mobile-Health application software) is implemented by combining structured query language (SQL) together with a high level language (Java) which allows the creation of user interfaces and database access in one application (Fig. 7).

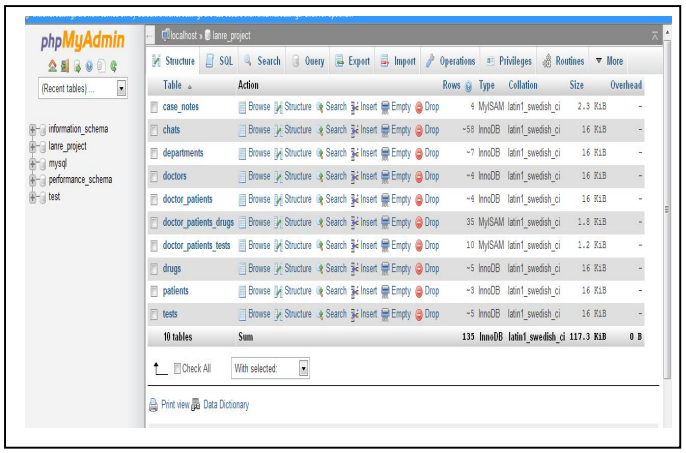


Figure 7. Database schema using mysql

XI. IMPLEMENTATION

Implementation is the stage in the project where the theoretical design is turned into a working system through the use of a programming language. In the implementation phase, the new system is developed, installed and operated. The system is designed and implemented such that the following are carried out during its use:

User validation: To be able to use the application, students are to register as patients into the mobile-Health application software and doctors, pharmacists, laboratory scientists are to register into the electronic-health application software with a username and password on the first login to both applications (Fig. 11).

Patients' Registration: Patients are to be registered on the system via mobile-Health application software on smart phones.

Doctors' Registration: Doctors are to be registered on the system via electronic-Health application software on desktop.

Pharmacist' Registration: Pharmacists are to be registered on the system via electronic-Health application software on desktop.

Laboratory scientists' Registration: Laboratory scientists are to be registered on the system via electronic-Health application software on desktop.

Usage: Doctors will login to the Crescent Health Information System to chat with patients and prescribe drugs (Fig. 13) and tests to the patients. The pharmacists will also login to Crescent Health Information System to update the drugs record and also chat with patients to verify drugs in the clinic. Laboratory scientist will login to the Crescent Health Information System to update tests record and chat with patients to verify information about a particular test. The patient will log into Crescent Mobile Health to chat with doctors, laboratory scientists and pharmacists and can view drugs and tests prescribed to them by the doctor. Users of Crescent Health Information System can update their profile on the application.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

XII. SCREEN OUTPUT FOR ELECTRONIC-HEALTH APPLICATION SOFTWARE

The screenshots below describes the output of the codes in PHP using sublime text and its interaction with MySQL database for Crescent Health Information System. The pages are described with respect to their workings.

A. *Login Screen*

This is the login page of the application. It will appear once the user runs the system. The user must enter the valid username and password to login before they can start using the application. A user without username and password will register by clicking register as a doctor on this screen (Fig. 8).

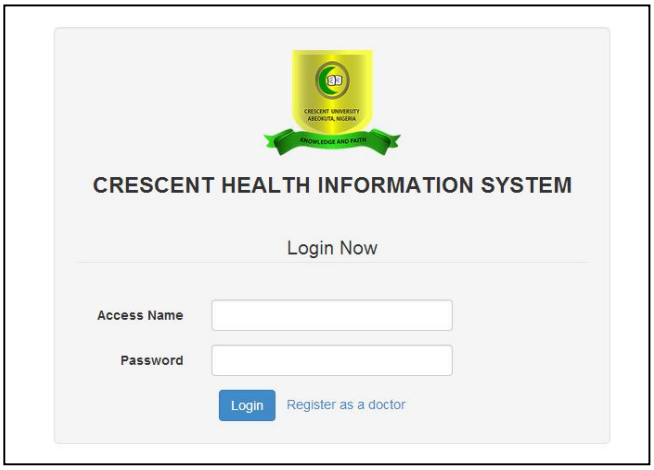


Figure 8. User login screen

B. *Chat and Prescribe Drugs Screen*

This screen allows one to view the chat section and prescribed drugs section where the doctor can prescribe drugs to a patient while chatting with the patient (Fig. 9).

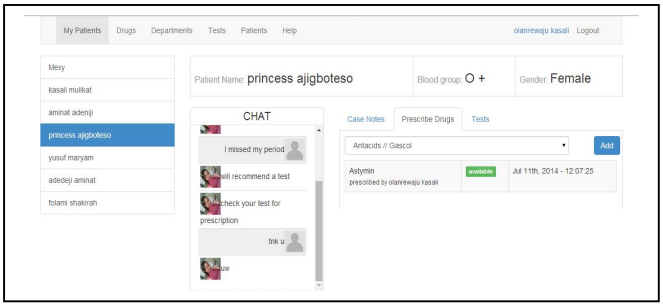


Figure 9. Chat and Prescribe Drugs Screen

C. *Patients Module*

Patients module is provides an over view of all patients that have registered into the mobile-Health application software. It consists of information such as; username, status, name, blood group, gender, date registered and actions (Fig. 10).

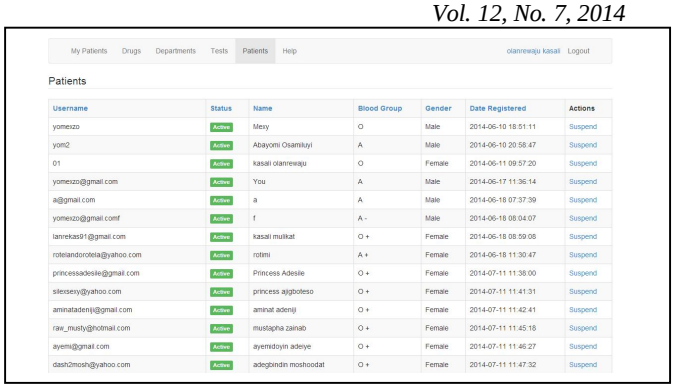


Figure 10. Patients module screen

XIII. SCREEN OUTPUT FOR MOBILE-HEALTH APPLICATION SOFTWARE

The screenshots below describes the output of the codes in java using android development studio IDE and its interaction with MySQL database for Crescent Mobile Health. The pages are described with respect to their workings.

A. *Sign In Screen*

The sign in screen is the first interface of the application where the user is required to enter email address and password to log into the application. A new user can register into the application by clicking “go to register” (Fig. 11).

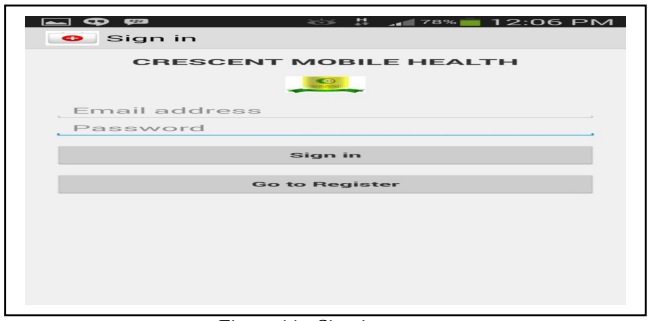


Figure 11. Sign in screen

B. *Chat Screen Displaying Sub Modules*

The image below displays a chat between a patient and a doctor where by d patient has clicked on the home button and a list of sub modules such as; my tests, my prescriptions, health tips and help file pops up (Fig. 12).

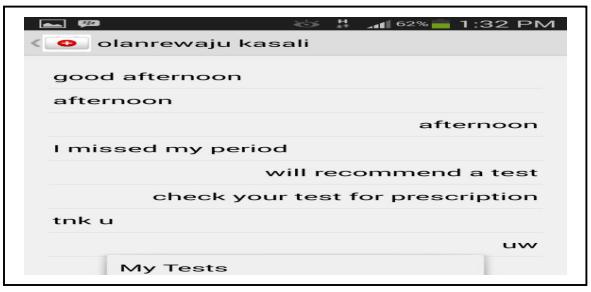


Figure 12. Chat screen displaying sub modules

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

C. My Prescriptions Screen

My prescriptions screen provides a view of the drugs the doctor has prescribed for the patient. It also consists of the name of the doctor that prescribed the drug, the date and time the drug was prescribed (Fig. 13).

Figure 13. My drugs screen

CONCLUSION

Crescent Mobile Health has been designed to meet its aims and objectives with the aid of Crescent Health Information System. The Crescent Mobile Health application is capable of storing, processing and retrieving information needed by its users with high degree of accuracy and speed. Instant messaging occurs between Crescent Mobile Health and Crescent Health Information System via pusher which provides an implementation of sockets that enables peer-to-peer communication.

Crescent Health Information System is capable of carrying out some functions such as registration and updating of records of doctors, pharmacists and laboratory scientists, drugs and tests inventory, prescription of drugs and tests to patients, viewing of online and registered patients, instant messaging between its users and patients. Crescent Mobile Health is also capable of carrying out functions such as registration of patients, viewing of drugs and tests that has been prescribed to a patient, viewing of health tips and help file, instant messaging between patients and doctors/pharmacist/laboratory scientist.

REFERENCES

[1] Cipresso, P., Serino S., Villani D., Repetto C., Selitti L., Albani G., Mauro A., Gaggioli A., Riva G. (2012). "Is your phone so smart to affect your states? An exploratory study based on psycho physiological measures". Neurocomputing 84: 23–30.

[2] Vital Wave Consulting (February 2009). mHealth for Development: "The Opportunity of Mobile Technology for Healthcare in the Developing World". United Nations Foundation, Vodafone Foundation. p. 9.

[3] Germanakos P.& Samaras G. "A Mobile Agent Approach for Ubiquitous and Personalized eHealth Information Systems" Proceedings of the Workshop on 'Personalization for e-Health' of the 10th International Conference on User Modeling (UM'05). Edinburgh, July 29, 2005, pp. 67–70.

[4] Ossman, J. (2010, May). "Barriers and Gaps Affecting m-Health in Low and Middle Income Countries: Policy White Paper". New York Center for Global Health and Economic Development, The Earth Institute, Columbia University.

[5] Hanauer D.A et al. (2009): Computerized automated reminder diabetes system (CARDS): e-mail and SMS cell phone text messaging reminders to support diabetes management.

[6] Istepanian, R., & Lacal, J.(2003). Emerging Mobile Communication Technologies for Health: Some imperative notes on m-Health. In IEEE (Ed.), The 25th Silver Anniversary International Conference of the IEEE Engineering in Medicine and Biology Society. Cancun Mexico; IEEE.

[7] Christine Zhenwei Qiang, Masatake Yamamichi, Vicky Hausman and Daniel Altman (2011). "Mobile Applications for the Health Sector", ICT Sector Unit World Bank December 2011.

[8] Blumberg. S, et al., "Wireless substitution: State-level estimates from the National Health Interview Survey", 2010-2011 in National Health Statistics Reports 2012, CDC.

[9] http://www.coyneclients.com/quest_diagnostics/gazelle_smpr/

[10] https://www.itriagehealth.com/what-is-itriage

[11] http://www.m.webmd.com/default.htm

[12] http://www.tutorialspoint.com/android/android_architecture.htm

[13] "Customer Information Manager (CIM)" . SOAP API Documentation.Authorize.Net. July 2013. Retrieved 2013-09-27.

AUTHORS PROFILE

Ayangbekun, Oluwafemi J received his Bachelor of Technology (BTech) in Computer Engineering from Ladoke Akintola University of Technology Ogbomoso, Nigeria in 2003. He also obtained his Masters of Science (MSc) in Computer Science from University of Ibadan, Nigeria in 2007. He is presently a PhD researcher in the Department of Information Systems, University of Capetown, South Africa

Kasali Olanrewaju M. received her Bachelor of Science degree in Information Technology from Crescent University Abeokuta Ogun State Nigeria in 2014.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

System Analysis and Design for integrated sponsored SMS/USSD Based M-Services

A case study of Maternal Health M-Service in Tanzania

Timothy Y. Wikedzi[1], Ramadhani S. Sinde[2]
Computational and Communication Sci & Eng
Nelson Mandela African Institution of Sci & Tech
Arusha, Tanzania

Dan K. McIntyre [1]
Information Technology
University of Iringa
Iringa, Tanzania

***Abstract*--Mobile phones have proven to be the best way of providing reliable access to information to people in low and mid income countries where other forms of communication perform poorly. As a result of the wide spread of mobile phones, there has been an increase in number of Mobile Application (M-Services) which are being used as a tool for disseminating different type information to people. M-Services of this nature are established to address informational challenges that are faced by people especially low income people. Because of this then, these projects must be sustained so that people can enjoy the benefits of it. Contrary to this, reports show that most of these M-Services are facing the challenge of cost of operating them, which in a direct way affects the sustainability of these services. In this paper therefore we present an analysis and later design of a noncommercial M-Service, which integrates advertising functionality as a tool for subsidizing the cost of operating M-Services. To achieve this we have employed some concepts of Information System Analysis and Design (ISAD) as the guiding principle towards achieving our design. A prototype of M-Health is used for the study.**

Keywords-M-Service; ISAD; Ad, USSD; SMS; Mobile; Sustainable; Cost of operation.

I. INTRODUCTION

The mobile communications technology has quickly become the world's most common way of transmitting voice, data, and services in the developing world. They carry a potential of being the best media for dissemination of information because mobile services are widely available and inexpensive, [1]. Mobile phones are less inhibited by traditional access barriers that hinder the widespread use of many other communications technologies including geography, socioeconomic status, infrastructure such as electricity and literacy, [2]

The potential of Mobile Technology being the right tool for disseminating information has led to establishment M-Service projects that are initiated to disseminate information to various social and economic groups within the society. An example of such projects operating in Tanzania and internationally are Mobile Alliance for Maternal Action (MAMA), M-Health Alliance, e-agriculture and Maji Matone and many other applications for Farmers, Sports, Educations etc. Existence of these projects have been reported to Improve livelihood of people by providing them with information which is an important tool for making informed decisions and in staying updated, [3]. [4] also reported that M-Services can contribute in fighting poverty by facilitating the convergence of local and global knowledge and disseminate it to the rural areas so as to improve economic production capacity in the settings in which the majority of the poor live.

Despite the fact that M-Service projects have proven to be a potential way for disseminating information to people, they are challenged by cost of service which is directly affecting sustainability of such services (MAMA 2013). mHealth Alliance report affirms to the fact that M-Services improves access to information, But the question of financial sustainability and ultimately "Who pays?" poses persistent challenges and barriers to scale and investment in such projects. Studies show that cost of information is among the major barriers to effective use of information which has resulted in underperformance and in some cases, total failure of M-Service e.g. Maji Matone Project in Tanzania. The experience of MAMA is a similar case for most other M-Services even those which are not related to M-Health. A report by mHealth Alliance on usage of mobile phone applications for disseminating information indicated that, at present, the majority of M-Services, particularly in low and middle-income countries, are dependent on donor funding. The report states further that, this model of financing is unsustainable because of the lack of certainty that funding will be

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

renewed and always the funds are limited to allow the project to run in full scale. As a result, most of these M-Service projects do not survive because of their dependence on this form of financing. The question, who pays? Is at the centre of most M-Services discussion.

According to [5] report, Securing sufficient revenue is still a challenge for most providers of noncommercial M-Services. [5] Presents a solution that, M-Service providers must develop a creative mix of revenue streams while taking into account the affordability of services. The report identifies the challenge of limited income for people in rural areas and so recommends an advertising model to be used to cover the cost of operating the M-Service.

There are M-Services which have succeeded and based on the current trend they have shown potential for further adoption (intermedia 2013), examples of such services are Mobile Money (M-Money) and Mobile Banking. Which uses a self-financing mechanism by charging users per transaction. There are also existing commercial M-Services agencies which run mobile services for information access such as Health information and agricultural information on weather and market prices for farmers as well as other forms of SMS campaigns, [6], [7]. Instead of providing free access to information these agents charge all their subscribers an access fee that ranges between 250Tsh, 500Tsh or higher for access. These models of financing M-Services promise sustainability but they pose a threat of Digital Divide between those who can pay for access and those who cannot pay. Cost of service in running M-Service has direct impact to the sustainability of M-Services. Donor funded M-Service project will only survive as long as the project gets support, but as the size of the project grows it then becomes a direct threat to sustainability of the M-Service. A sustainable financial model is needed to ensure survival of M-Service projects [8]

This study is set to adopt the recommendation from [5] and design an M-Service that integrates Advertisements functionality as system module for subsiding cost of operating an M-Service. Advertisements have widely been used in the field of Information technology as a mechanism for covering cost and of course making profits. Companies like Google, for example, make a fortune through advertisements. The potential of advertisements in covering the cost of access has benefited a lot of other companies like Facebook and Yahoo which provide users with free access at the cost of receiving advertisements. Nowadays there are thousands of free mobile apps in the market which use advertisements as a tool for generating income. However in this study we are set to present an idea of SMS advertisesments as a potential solution for free access to information for all categories of mobile phone users.

Implementing a sustainable M-Services is an important aspect towards ensuring long term benefits of those projects. In this paper we present the Analysis and Design of an Integrated M-Service that addresses the challenge of cost of operation by using SMS advertisement support as an integral part of the system. A case study of M-Service for Maternal Heath in Tanzania is used.

II. LITERATURE REVIEW

A. Mobile Application as M-Service

[5] Defines mobile Applications as software designed to take advantage of mobile technology. In this paper we specifically refer to Mobile Applications as SMS Based Mobile Application (M-Services). This way we put a distinction of M-Service from more complex and advanced Mobile Applications that run on smartphones. The reason to do this is that, although mobile phones are widely spread in low and mid income countries, the majority of people, especially in low income countries, own basic phones which have limited features compared to smartphones.

B. Sustainable M-Service

[9] Defines Sustainability as the process of maintaining something that already exists over time without needing an outside support for it to continue existing. Whereas [10] provides us with a more specific definition of sustainable IT as a Technology that is capable of being maintained over a long span of time without being affected by the changes in both hardware and software. This definition presents to us two important things to consider when talking about sustainable IT, in this case an M-Service, one is the life span of the system, and two is the ability of a system not being affected by the changes in hardware and the software.

C. SMS

Short Message Service (SMS) is a text messaging service component of phone, Web, or mobile communication systems. It uses standardized communications protocols to allow fixed line or mobile phone devices to exchange short text messages, [11]. Content of one SMS is limited to a maximum160 ASCII characters. SMS are sent to mobile phones via the SMS Gateway. Using SMS technology, it is also possible to send them in Bulk, in

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

which one or more SMS are sent to more than one user. This offers a better way of reaching out to more people at a time.

D. Ad

This is a term that is used to refer to Advertisement. In this context, it always refers to a text base advertisement that is intended to be sent to the end user inform of an SMS.

E. SMS advertising

SMS advertising is a subset of Mobile advertising, a form of advertising via mobile phones or other mobile devices. Other forms of mobile advertising which are specifically common in smart mobile devices are Mobile Web Banner (top of page) and Mobile Web Poster [12]. The focus of this paper though is on SMS advertising, which has been reported to be the leading form of Mobile advertising worldwide, for the reason that our proposed system is also for no smartphone users.

F. Approaches of pushing SMS advertisements

1) *Receiver Opt-in:* This form of SMS advertisement involves requesting for the consent of the potential receiver before starting pushing of advertisements.

2) *Purchase of Receiver numbers from third part source:* This form of advertisement involves purchasing of receivers contacts from third party companies and using them to push SMS advertisements

G. USSD

[13] Defines *Unstructured Supplementary Service Data (USSD)* as a communication protocol used to send text messages between a mobile phone and applications running on the network. It is a messaging service used in Global System for Mobile Communications (GSM) networks similar to SMS. However, unlike SMS; USSD provides session-based connections. Because of its real-time and instant messaging service capability, USSD service offers better performance and is much cheaper than SMS for two-way transactions. This service is unique and only available to GSM networks. The following are the advantages of USSD as described by [13].

1) USSD code format

USSD communication is initiated by dialing a special code. USSD codes comprise of asterisk (*), followed by a combination of digits (0 to 9) and a hash (#) Example *150*00#. The * and # codes are used to signify the beginning and end of the request.

2) USSD Architecture

According to [13] The USSD architecture basically comprises

- The network part that includes the Home Location Register (HLR), Visitor Location Register (VLR), and MSC
- Simple Messaging Peer-Peer (SMPP) interface for applications to enable services
- USSD Gateway and all specific USSD application servers

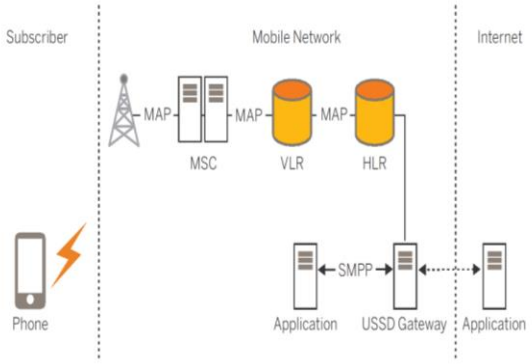


Figure 1: Elements of the USSD Mobile Network (Source Aricent)

While USSD seems to be a complex architecture, the focus of this study is not in analyzing USSD architecture but rather to make use of it as a black box. Our only concern is on the Internet and Subscriber sides. In this paper we present the analysis and design of an Application that that interfaces with USSD to reach the end users.

III. OVERVIEW OF THE SYSTEM

In this paper, we present the analysis and Design of an M-Service (m-Health) that rely on sponsorship (advertisement) to convey maternal health information to the public of Tanzania. The proposed system will involve the use of mobile phones to deliver messages to the target audience. The system will allow the user to subscribe and get information via short messages. All messages stored on the system will be provided by health professionals to ensure that the information sent to users are correct. These health professionals will be assigned special accounts that they will use to access and work in the system. End users will be able to retrieve this information via SMS using their mobile phones. They will also have the opportunity to send various questions and receive answers from medical professionals. The system is designed with a capability of sending advertisements to end users as a way of covering the cost of operation. To ensure future sustainability of this system, the system design also

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

integrates a payment system for economic buyers who will not want to receive advertisements. Payment service will also be useful in time if there are no funds to cover the cost of free access and in meeting the needs of those who would want unlimited access. However, the focus of our analysis and design is on Sponsored access (access through advertisements).

A. *Integrating Technologies*

The Architectural design of our M-Health is based on four tiers architecture (see **Error! Reference source not found.**) as it has been proposed by [14]. They identified that using a four tier architecture for mobile applications offers more abstraction, more independence between components of the system, and hence a more flexible way of implementing mobile applications. We have adopted this architecture because we see it fitting our design idea. It also gives us a great opportunity to plan for growth as mobile technologies evolve so fast and the future holds that more mobile phone users will be upgrading to smartphones and other devices.

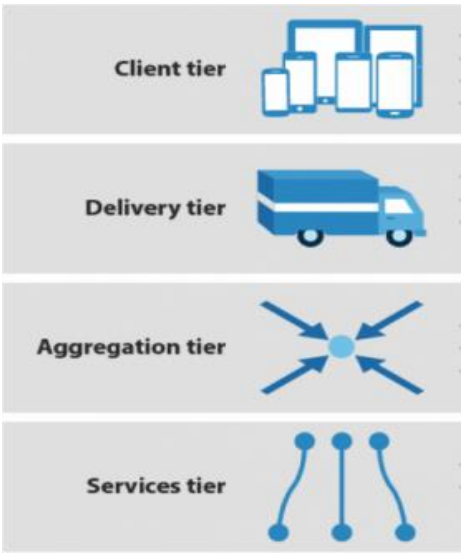


Figure 2: Adopted Four Tier Architecture (Source: http://blogs.forrester.com/)

B. *Implementing the Four Tier architecture for our M-Health*

- *A client tier*: Represents end users means of accessing our M-Health System. The client tier in this case refers to both mobile phones and web clients. This presentation layer creates a separation between the applications and accessing devices and the backend services that the application delivers.
- *Delivery tier*: Is the channeling tier. It handles the physical routing of requests and information to and from the system. This is the direct interface of the systems and the end users. In our M-Health Systems this represents Telecommunication channels and services such as USSD and SMS. It also represents the Internet as the channel through which Medical Experts will be interacting with the system.
- *Aggregation tier*: Acts as a bridge between data and upper tiers. It provides means for handling all data requests and exchanging data between service tier and the Delivery tier. Our M-Health System will run on PHP as the language for our application.
- *A services tier*: Handles all data functionalities. It includes the database that stores and manages all data. The choice of Database for our M-Health system is a MySQL database. The reason for choosing this type of Database management system is that it's open source and has a strong community for support.

C. *Sponsored (Advertisement) access Request life cycle*

Under this subsection we describe the concept of the advertisement mechanism for our proposed M-Service. Of the two approaches for pushing advertisements to the end user, we will be using the opt-in approach. Therefore, all our designs are based on an assumption that the end user opted to receive an advertisement

1) Information Requesting

Information processing begins with the user, sending a USSD request from his mobile phone. After dialing a special USSD code, the user receives a USSD menu from which he can choose categories of information to access. Information request is limited to categories after which a random information will be sent to the end user base on the selected category.

USSD menu generation

The USSD menu is dynamically generated from the defined categories of information in the database. This menu gives a user access to various types of information that is stored in the database. Generating menus dynamically adds more flexibility to the system.

2) Advertisement processing

a. Fetching of an Ad from the database

All advertisements will be stored in the database and associated with sponsors who will be registered into the system. The process of fetching an ad from the database is triggered internally after the M-Service has received an access request. Before an ad is sent, the system checks to see if there are any existing ads. The conditions for an ad to exist is that there should be at least one sponsor with

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

some remaining balance. This is how the system will know that a message can be sent. Once the ad has been confirmed to exist the system will proceed on fetching the current ad in the queue and submit it for further processing

b. Processing ads based on available sponsors.

In this initial design, all ads will be sent in a queue that will rotate around all available sponsors whose balance is above zero. The future design of this system will allow ads to be sent to a client base on the relevance of the locality as the first priority. This way the system will be able to send context based and relevant ads to end users. This of course will be achieved with an upgrade to smartphones.

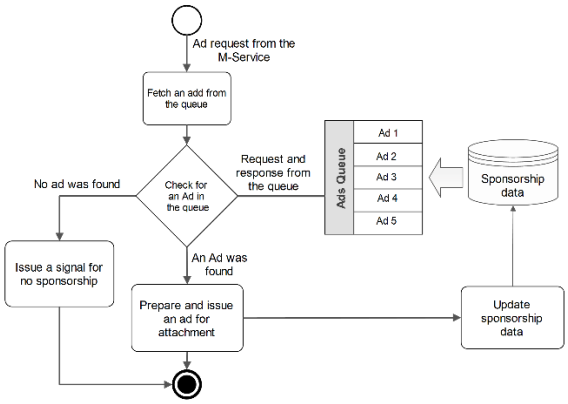


Figure 3: Design concept of the advertising mechanism

3) Fetching of information

Once sponsorship has been confirmed, content is then prepared to be sent to the users. As it has been stated earlier, information to be sent to the user will depend on the category of information that was picked by the user.

4) Integrating an ad with requested information

The next step after fetching both the ad and the information is to bind them together and send them as a single data packet to the user. There are two alternatives on how this information will be sent to the user; first, we send two separate SMS one for the ad and the other with the requested information. Second, we send one SMS containing both the ad and the content.

5) Content Delivery

Though all information requests will be issued via the USSD service, the requested information will be sent back to the user via an SMS. This is because USSD messages are session based, so once the session is over the returned message is lost. Also USSD messages cannot be saved in the inbox of the end user's phone.

IV. ISAD FOR SUSTAINABLE M-SERVICE

In Information Systems (IS) development, Analysis and Designing is used as a tool for developing successful IS. History and experience prove that most of the IS projects that had in one way or another ignored ISAD suffered great failures in the form of exceeding budgets, failure to meet the running cost of the system and in some cases total rejection of the system by the users. ISAD as a tool is a part of System Development Life Cycle (SDLC). Dennis, A. et el, 2012 define the Systems development Lifecycle (SDLC) as the process of determining how an information system (IS) can support business needs, designing the system, building it, and delivering it to users. The key concept in this definition is that IS are need based, meaning that there is no purposeless IS as they are all developed for the sake of addressing a certain need. In this paper, we present ISAD considerations for sustainable M-Service. We present the analysis and Design of an M-Service that integrated Advertisement as a tool for subsidizing the cost.

A. Identification of end users

As a part of ISAD, the project must clearly identify the end user, in the case of M-Service the end user is referred to as the individual(s) that and are intended to be accessing information offered by the service. The process of identifying users of an M-Service is highly dependent on the type of M-Service. If for example a service is meant for Farming activities, then clearly it means that farmers form a primary group of end users of that service. In the case of an M-Health (Maternal Healthy) system which we refer to as a case study in this paper, the end users of the system are primarily Pregnant women, Fathers, Mothers and all those who would benefit from accessing maternal health information. User identification as an aspect of ISAD is necessary to ensure that a solution based M-Service is developed and better design is later achieved.

B. Considering the cost of operation

Well, this is the focus of this study. As it has been reported that covering the cost of operation is a challenge that most M-Services face,. The cost of operation is a critical factor to be considered in any M-Service since they all rely on Mobile communications which are run commercially. Mobile network operators dominate the M-Services ecosystem in developing countries. They serve as gatekeepers, and dictating matters related to revenues including requiring payments for the service [5].

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

M-Services are categorized into two major types; commercial M-Services which are typically developed to deliver information or conduct transactions (or both) and noncommercial M-Services which are developed only to provide information [5]. Commercial M-Services require end users to pass through some form of payment when accessing the service. Sponsored (noncommercial) M-Services on the other hand have a third person who covers the cost so that end users can enjoy a free service. Running a Commercial M-service is clear when it comes to costs of operating them. Because end users are charged for the service and in so doing the host of such application makes a profit and covers the cost of operations. Examples of such M-Services are M-Banking, M-Money etc. In Tanzania the leading Commercial M-Service provider is Vodacom with their M-Pesa service which is a form of M-Money service that offers a service of mobile money transactions.

The Analysis and Design focus of this paper is on Sponsored M-services because, they have potential of or extending information accessibility to people of low income. Considering the current economic state of people in rural areas of Tanzania which according to census results of 2012 more than 60% of the whole population live in rural area where there is extreme poverty.

C. System and User interactions

The question of how users will interact with the system is critical in any IS project because it determines the type of technologies to be used as well as in planning for the cost of operation. Our M-Heath System offers two types of users Web Client users (Administrator and the Medical Experts) and Mobile subscribers who are the end users of our M-Health System. Web users interact with the system through Web browsers and internet whereas all our mobile subscribers uses their mobile phones to access the system. The mode of communication between the System and the subscribers will be via SMS/USSD services. Subscribers will also be able to post questions by sending them via SMS to a special number. Administrator and Medical Experts will be able to add new content and manage the system from the backend.

D. Requirements gathering and Analysis

[15] Defines Requirements as statements of what the system must do or what characteristics it needs to have. Any IS relies on requirements for its implementation an M-Service is no exception. However M-services present a challenge when it comes to gathering of requirements, M-Services are IS that are not domain based. They are developed to be used as a medium of disseminating information to open end users, it could be a well-defined group of end users, like in this case pregnant women for example. But still these people are not organized under one domain, they are all independent and with freedom to choose to adopt the service or to remain as watchers. We understand though, the idea of our M-Service is that it will be delivering simple text messages either in form of SMS service or USSD messages. End users of our application need not to know how the backend of our application is structured. Our Analysis and Design is therefore characterized by the idea of a information system that delivers information to end users in a manner that users can afford to get it, access it easily, and understand the content.

1) M-Health application requirement

Since the focus of this paper has been on Sponsored M-Services we will then discuss requirements based on that sector. Before defining requirements we need to describe how it will work with use cases. In a domain limited IS projects, requirements are derived from the business processes of a particular domain e.g. Human Resource IS, Academic IS, Payroll IS etc.IN the other hand, M-Services are open ended systems, meaning that the developer proposes the requirement and implements the system. The role of end users is not to provide initial requirements but rather to give feedback that will later improve their experience with the service.

2) Functional Requirements

Functional requirements are those requirements that are used to illustrate the internal working nature of the system. They describe what tasks the system should perform.

a) Subscribers

- Subscribes
- Send questions
- Access Information via mobile phone

b) Administrator

- Manage Doctors
- Monitor Subscribers
- Manage Sponsor Information
- Access Information

3) Nonfunctional Requirements

a) Operational

- USSD/SMS service

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

- The system should interface with SQL database.

b) Maintainability & Upgrades

- System should allow upgrade to smartphone usage
- System should allow updates and upgrades without affecting users' experience

c) Acceptance

- System Language should be Swahili

These are a summary of all requirements that as a developing side we have come up with. It is of course true that some or most of these requirements will change overtime as users will require some aspect of the systems to be improved, added, or dropped. There is at all-times a call for better future versions of the system and services.

E. Structuring of the Requirement

Having defined the requirement for our M-Service we move to the next stage of systems analysis. According to [16] system analysis phase consists of two parts, determining requirements and structuring requirements. During structuring of the requirements the goal is to interpret and model processes and data for our application. A common tool used for structuring information is the Data-Flow Diagram (DFD). [16] Describes **data flow** as data that are in motion and moving as a unit from one place in a system to another. DFDs use four symbols, data flows, data stores (e.g. database), processes, and sources/sinks (or external entities) to represent both physical and logical information systems.

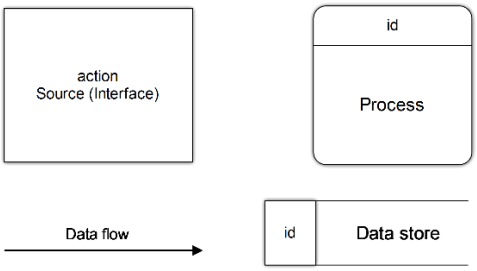


Figure 4 : DFS tools

1) *Process flow modelling:* Process modelling is meant to provide a description of how data flows from one point to another. Our M-Health system has three possible users who will act as the action source to our system. These users interact with the system with different roles and each causes different types of data to flow.

a) Context Diagram: A Context Diagram in this case is meant to give a general concept of the system and data that will be flowing as ads to the identified interfaces (users). Figure 5 Shows the context diagram for the proposed System. It consists of the main process Health SBMA (SMS Based Mobile Application)

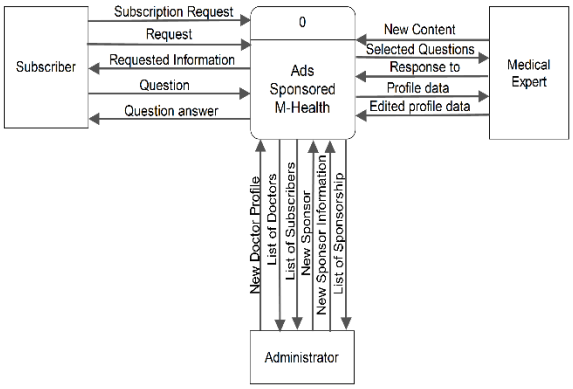


Figure 5: Context Diagram showing general data flows for our mHealth

b) Level 0 Process Diagram

Expounding the context diagram, we create the object Level 0 diagram. The level 0 diagram basically breaks down the Health SBMA (SMS Based Mobile Application) as identified in the context diagram. Figure 6 show the diagram for level 0 of the proposed M-Health System.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

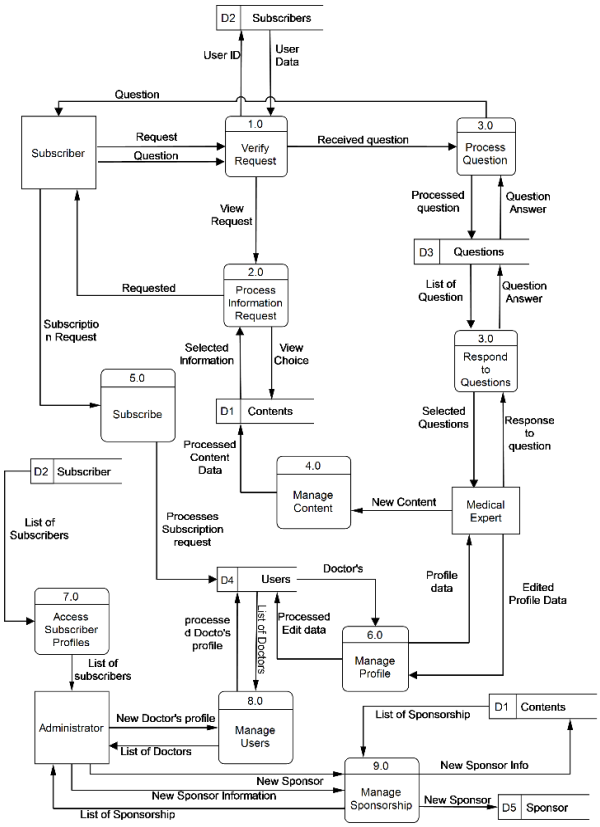


Figure 6: Level 0 data flow diagram describing processes in more details

c) Level 1 Process Diagram

Level 1 diagram for our M-Health describes the breaking down of process 2.0 which handles the task of processing information requests. There are two types of request that the system can receive, one is of free access to information and the other one is for the paid access.

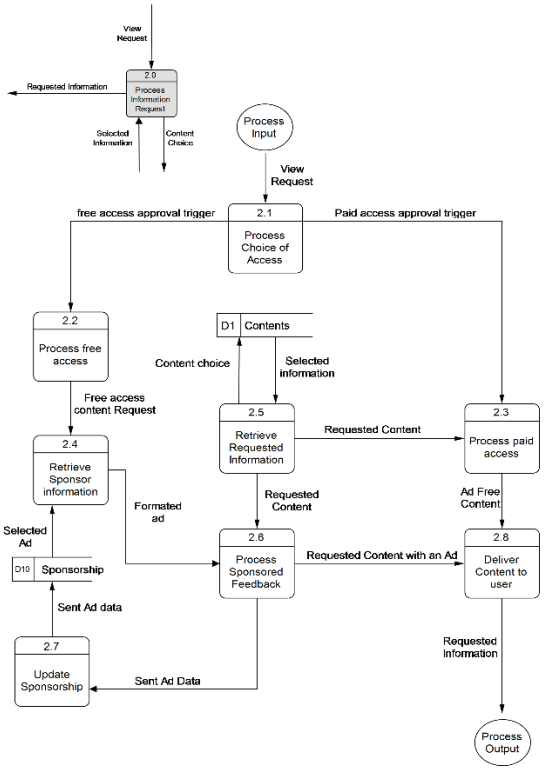


Figure 7: Level 1 Data low diagram for Process 2.0 in Level 0

2) Conceptual Data Modelling

[16] describes conceptual data modelling as a way of representing organizational (Information Systems) data. The goal of data modelling is to show as many rules about the meaning and interrelationships among data as possible, independent of any database management system (e,g MySQL, Oracle, MS SQL, SQLite). The common tool that is used to model the data are Entity-relationship (E-R) data models. These are diagrams that show how data are organized in an information system. ERD uses a special notation of Rectangles (to describe entities), diamonds (to describe relations), and lines to represent as much meaning about data as possible. The deliverable from the conceptual data-modeling step within the analysis phase is an ERD Diagram

During the analysis of requirements for our M-Health we had put focus on data to gain the perspective on data needed to develop a data model. The following entities have been identified; (1) USER_GROUP (Defines roles of users e.g Doctors, Administrator), (2) USER (Handles users' profile data), (3) SUBSCRIBER (Stores subscribers information) , (4) QUESTION (Stores posted questions), (5) ANSWER (stores responses to questions), (6) CATEGORY (Stores categories of information which forms menus and submenus), (7) CONTENT (Stores

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

information that subscribers can access), (8) RECEIVED_SMS, (9) SPONSOR (Stores sponsors information), (10) ADS (Stores all the advertisements)

The conceptual data model of these schemas is represented diagrammatically by Figure *8*

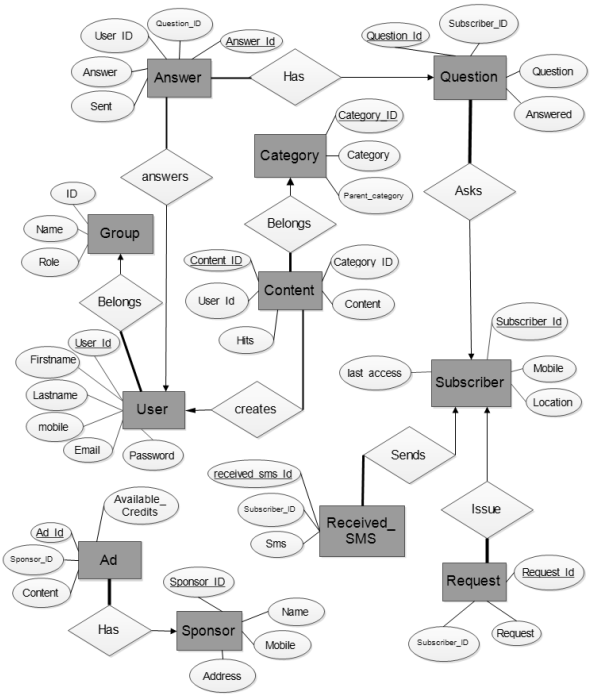


Figure 8: ERD Diagrams describing the conceptual model of our mHealth System

The ERD diagrams presented in this subsection forms the basis for the database implementation.

F. System Designing Considerations

1) Keeping it simple

SMS based application are characterized by the size of the SMS which is determined by a fixed number of text characters. The size of SMS directly affects the cost of operation because the cost of sending SMS increases with the number of Messages. For this reason, then it is important that the overall application is kept simple and more focus to be put on means for optimizing the amount of information for an SMS.

2) *Information Design*

[17] States that a high quality information design communicates information in a manner appropriate and pertinent to a reader's situational context. It must focus on the reader ability to understand it and to extract meaning from that information. Information Design is one of the very important factors for a sustainable M-Service. Our M-Health System is meant for Tanzanians and so information design aspect should consider the nature of these people to determine the type of information and in which form that information going to be delivered. The national Language of Tanzania is Kiswahili? And it is both the official language and the language that most people understand. It is obvious then, one of information design criterial for our M-Service is that its content must be in Swahili. Another very important factor to consider is, as has been mentioned before, that this application relies on SMS/USSD services. The cost of these services is determined by the size of message

3) Designing Interactive menus for USSD

All end users (Subscribers) of the system will access the system via their mobile phones. The system is designed to work for all types GSM of phones. **Error! Reference source not found.** Shows a summary of user actions when accessing our M-Health application via the mobile phones. The flow of actions is numbered from 1 to 12. (1) User enters the USSD code to access information in this case it is *31022, (2) If the user is accessing the system for the time he will be asked to subscribe by sending a messages to a special number (registration could free or onetime payment of 250Tsh). (3) Once in the system the user will receive a USSD message informing him that the service is free but he will receive an ad. (4) User can then choose to proceed or quit. (5) Upon agreeing to continue a menu of information categories will be presented.(7) More submenus will be presented for user to narrow down his choice. (9) After reaching to the last subcategory, user will receive a USSD SMS informing him that information will be sent to his phone as an SMS. (10) User will then receive an advertisement with a code that he will have to send to the system via USSD. Sending of this code is a way of making sure that at least the user read the Ad SMS. (12) Lastly the requested information

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Figure 9: Action Flow UI's whowing end Users Usage of the SYstem (Content on phone is in Swahili because the targeted end users are the Tanzanians)

4) Web interface Design

Both system administrator and the doctors will access the systems by using a web interface. This way they will be able to add and manage new users and manage new content respectively.

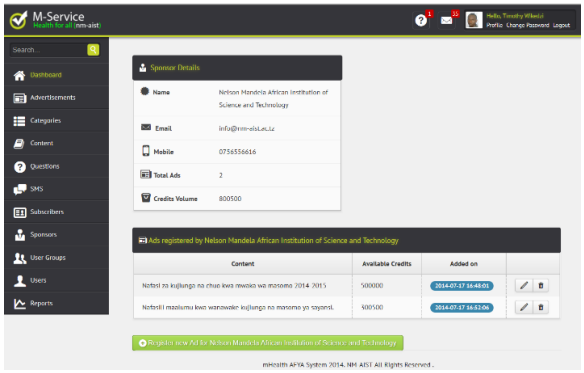


Figure 10: Sample web Client page interface for administrators and Medical Experts to access the system.

G. Ads Consideration

In the context of this paper, we strictly consider advertisements as an information and therefore the concepts of information design also apply to them. It is important for all the advertisements to be categorised so that right ads can be delivered to the right people. Ads can be delivered base on Locality of an individual but also base on the type of content that the user requested. Some ads may be useless to people who cannot afford to pay for what is advertised, Though not each of our subscribers will be poor, it is still important to determine the type of advertisments to be sent to subscribers.

H. Defining the right source of information

Sustainable M-Service must have a reliable source of information. Reliability of information in this case does not mean only availability of such information, but it also implies trusted sources of right information. We do not want to feed our users with false information this could easily lead to misleading and loosing trust of our users. Our M-Health service will rely on Qualified Doctors and other Medical practitioners as the source of content for our M-Health.

I. Conclusion

In this study, we set out to do a case study on how to develop a system which is maintainable and sustainable in a community. By reflecting on the process, and the system being designed we intended to draw conclusions of the important aspects which exists when developing such a system. Our study is based on a M-Health prototype application that was analyzed and designed based on the knowledge of how existing M-Services work from end users perspective. This way we were able to design our M-Health prototype which will later be improved through feedback from end users.

We acknowledge the fact that a sustainable IS puts into consideration the growth aspect. It is then meaningless for

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

a system to be developed to offer a useful service or solution and due to failure of shifting to new technology. This being the case, we recommend that a sustainable IS should not be tightly coupled with technology but rather should run independently of it and be flexible to change and adopt new technologies whenever a change is required, example changing to smartphones and modern technologies that are supported by mobile devices. SMS Based M-Services must therefore be developed strategically ready to be upgraded into smartphone applications, this will guarantee greater chances of continuing provision of the service. We have chosen to adopt the four tier architecture that decouples Data, processing application, Distribution Channels and Client technologies. This way our application can integrate and work well with Mobile devices, Web Clients, Any Data source and can potentially be distributed via internet and GSM networks.

As an outcome of this study and analysis, we finally conclude by acknowledging that M-Services have proven to possess a great potential of addressing information challenges that are facing low and mid income countries, therefore they must be kept sustainable to ensure that the intended service is continually offered for the benefit of people.

RFERENCES

[1] J. Muthee and N. Mhando, "African Media Development Initiative Tanzania.," 2006.

[2] R. Genuchten, W. Haring, D. Kassel, and K. Yakubi, "Mobile phone use in Tanzania," 2012.

[3] A. S. Sife, E. Kiondo, and J. G. Lyimo, "Contribution Of Mobile Phones To Rural Livelihoods And Poverty Reduction Inmorogoro Region, Tanzania," EJISDC, vol. 42, pp. 1-15, 2010.

[4] A. K. Hassan and D. Semkwiju, "The Role of Mobile Phones on Sustainable Livelihood," The Economic and Social Research Foundation (ESRF), 2011.

[5] C. Z. Qiang, S. C. Kuek, A. Dymond, and S. Esselaar, "MobileApplications_for_ARD," 2011.

[6] BongoLIVE. (2014, July 8 2014). Targeted SMS Marketing. Available: http://www.bongolive.co.tz/sms_advertising.php

[7] PushSMS. (2014, July 8 2014). Mobile advertising. Available: http://www.push.co.tz/marketing/advertising/

[8] MAMA, "healthy pregnancy, healthy baby text messaging service Tanzania," 2013.

[9] J. Reynolds and W. Stinson, "Sustainability analysis, Primary Health Care Management Advancement Programme," 1993.

[10] G. Misund and J. Høiberg, "Sustainable Information Technology for Global Sustainability, Digital Earth," Information Resources for Global Sustainability Symposium, vol. 9, pp. 21.-25, 2003.

[11] Wikipedia. (2008). Short Message Service. Available: http://en.wikipedia.org/wiki/SMS#cite_note-1

[12] Wikipeda. (2014, July 8, 2014,). Mobile Avdertisment. Available: http://en.wikipedia.org/wiki/Mobile_advertising

[13] J. Sanganagouda. (2013). USSD: a communication technology to potentially oust SMS dependency.

[14] M. Facemire, J. McCarthy, and T. Schadler. (2013). Mobile Needs A Four-Tier Engagement Platform. Available: http://blogs.forrester.com/print/ted_schadler/13-11-20-mobile_needs_a_four_tier_engagement_platform

[15] A. Dennis, B. H. Wixom, and R. M., "Systems analysis and design," 5 ed: John Wiley & Sons, Inc., 2012, p. 104.

[16] J. S. Valacich, J. F. George, and J. A. Hoffer, Essentials of Systems Analysis and Design, 5 ed. New York: Pearson, 2012.

[17] M. J. Albers and M. B. Mazur, Content and Complexity: Information Design in Technical Communication: Taylor & Francis, 2014.

AUTHORS PROFILES

TImothy Wikedzi, is a Master's student at the Nelson Mandela African Institution of Science and Technology. He is pursuing a master's degree in Communication Science and Technology. He currently lives and study in Tanzania

Dan K. McIntyre, is an Adjunct Professor at University of Iringa in Tanzania. He is a an expert in Computer Science particularly Software Engineering. He currently lives in United States of America.

Ramadhani S. Sinde is an Assistant lecture at Nelson Mandela African Institution of Science and Technology. He is an expert in Telecommunication.

.

CONVERSION OF AN SR-FLIP FLOP TO A JK-FLIP FLOP

By

Prof. Olawale J. Omotosho
Babcock University,
Computer Science Department
Ilishan-Remo, Ogun State, Nigeria
Tel: +2348034951089

Engr. Samson O. Ogunlere
Babcock University,
Computer Science Department
Ilishan-Remo, Ogun State, Nigeria
Tel: +2348067622845

Abstract

This paper presents a design method to convert a conventional SR-Flip Flop to perform the functions of a corresponding conventional JK-Flip Flop. This requirement becomes very necessary because of the many applications of JK-Flip Flops in digital systems, especially in those systems that drive production industries. In such industries, uninterrupted production is one of the targets required to pay attention to in order not to lose production and consequently revenue. Equipment failure can be responsible for such an unwanted state of production. Therefore, downtime of any equipment becomes very crucial in the assurance procedure of associated equipment and instrumentation of a manufacturing plant. The cause of a large downtime of any equipment is mainly due to unavailability of spare parts and sometimes incompetence and inexperience of the Technologists responsible for the up-keep and assurance of these equipment and instrumentation. Technologist must be versatile in providing alternative solutions to existing provisions that is adequate to solve any prevailing situation which requires urgent attention to keep production going. Such experience is not only borne out of hands-on practice but can be acquired by sound theoretical knowledge of what to do. This paper examines a situation where a device (JK-Flip Flop) is not available to replace a defective one but an SR-Flip flop is configured to be used for the same purpose without degradation of performance.

Key words*—Conventional Flip-Flops, uninterrupted production, unavailability of spare parts, incompetence and inexperience of the Technologists, downtime of equipment, K-maps.*

Introduction

In industries, uninterrupted production is one of the targets of any manufacturing concern. Equipment failure can be responsible for such an unwanted state of production. Therefore, downtime of any equipment becomes very crucial in the assurance procedure of associated equipment and instrumentation of a manufacturing plant. The cause of a large downtime of any equipment is mainly due to unavailability of spare parts and sometimes incompetence and inexperience of the Technologists responsible for the up-keep and assurance of these equipment and instrumentation. Technologist must be versatile in providing alternative solutions to existing provisions that is adequate to solve any prevailing situation which requires urgent attention to keep production going. Such experience is not only borne out of hands-on practice but can be acquired by sound theoretical knowledge of what to do. This paper examines a situation where a device (JK-Flip Flop) is not available to replace a defective one but an SR-Flip Flop is configured to be used for the same purpose without degradation of performance.

1. A Conventional NOR-gate-SR-Flip Flop

The Truth Table of an SR-Flip Flop is presented on Table 2.1 using a NOR-gate configuration and the corresponding K-Maps from where appropriate minimal logic equations are derived.

TABLE 2.1: TRUTH TABLE & K-MAPS

S/N	SR-FF TRUTH TABLE		(a) K-MAP of Q_{n+1}		(b) K-MAP of Q_{n+1}

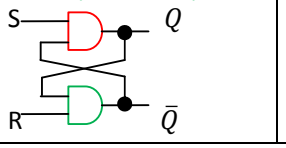
http://sites.google.com/site/ijcsis/
ISSN 1947-5500

	S	R	Q_n	Q_{n+1}
0	0	0	0	0
1	0	0	1	1
2	0	1	0	0
3	0	1	1	0
4	1	0	0	1
5	1	0	1	1
6	1	1	0	X
7	1	1	1	X

Q_n \ SR	00	01	11	10
0	0	0	x	1
1	1	0	x	1

$Q_{n+1} = S\bar{Q}_n$

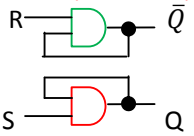
$\bar{Q}_{n+1} = RQ_n$



Q_n \ SR	00	01	11	10
0	0	0	x	1
1	1	0	x	1

$\bar{Q}_{n+1} = R\bar{Q}_n$

$Q_{n+1} = SQ_n$



In the K-Maps of Table 2.1 (a) & (b), we are concerned about the forbidden states in the lower portion of the truth table marked in black. There are two ways of combining the states as shown in K-Maps (a) & (b). These two ways are employed to convert the SR-Flip Flop into a JK-Flip Flop.

2. Conversion of an SR-Flip Flop to a JK-Flip Flop Using Table 3.1

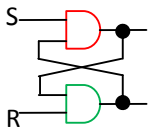
Table 3.1: Four Possible Logic Equations from K-Map Considering Only Forbidden States

(a) K-MAP of Q_{n+1}

Q_n \ SR	00	01	11	10
0	0	0	x	1
1	1	0	x	1

$Q_{n+1} = S\bar{Q}_n = S_1 .. (3.1)$

$\bar{Q}_{n+1} = RQ_n = R_1 \ldots (3.2)$

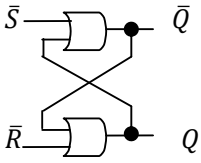


Right

OR-Gate

$\bar{Q}_{n+1} = \bar{S} + Q_n = S_1 .. (3.3)$

$\bar{Q}_{n+1} = \bar{R} + \bar{Q}_n = R_1 .. (3.4)$

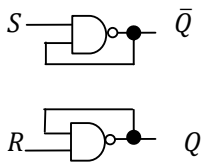


Wrong

NAND-Gate

$\bar{Q}_{n+1} = \overline{S\bar{Q}_n} = S_1 \ldots (3.5)$

$\bar{Q}_{n+1} = \overline{RQ_n} = R_1 \ldots (3.5)$

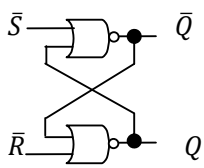


Wrong

NOR-Gate

$\bar{Q}_{n+1} = \overline{\bar{S} + Q_n} = S_1 \ldots (3.7)$

$\bar{Q}_{n+1} = \overline{\bar{R} + \bar{Q}_n} = R_1 \ldots (3.8)$



Right

This is applicable to NOR-gate SR-Flip Flop configuration. That is, AND-gated SR-Flip Flop and NOR-gated SR-Flip Flop will perform the functions of a conventional JK-Flip Flop (see prove in section 5).

Since the outputs (Q and $\bar{Q}$) of an SR-FF are complementary to each other, its SR inputs will never be confronted with S = R = 1 if it is triggered by its outputs. Hence, equations (3.1) & (3.2) are used to trigger an SR-FF at the instance when S = R = 1. That is, the SR-FF is triggered during its Forbidden states by its outputs as shown in Fig 3.1.

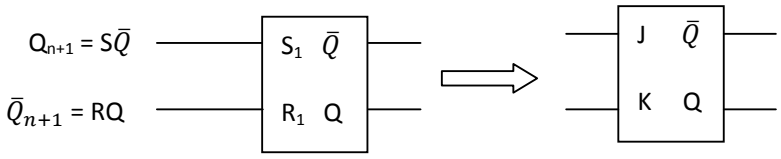


Fig 3.1: Modified SR-Flip Flop ⍰ JK-Flip Flop

Where,

'S' is a new Set-terminal of a modified SR-FF that will perform like a JK-FF where the Forbidden states of an SR-Flip Flop are converted into Toggle states. That is, S = J.

'R' is a new Reset-terminal of a modified SR-FF that will perform like a JK-FF where the Forbidden states of an SR-Flip Flop are converted into Toggle states. That is, R = K.

'S_1' is an old Set-terminal of an unmodified SR-FF that maintains its Forbidden states.

'R_1' is an old Set-terminal of an unmodified SR-FF that maintains its Forbidden states.

Four different cases/options will be considered using the modified SR-Flip Flop to obtain which of them will behave like a JK-Flip Flop as follows:

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Option-1 (AND-gate Configuration)

Equations (3.1) & (3.2) are already in AND-gate forms. That is,

$Q_{n+1} = S\bar{Q}_n = S_1$…….. (3.1) $\bar{Q}_{n+1} = RQ_n = R_1$……. (3.2)

Hence, the SR-FF can be gated as per these equations as shown in Figure 3.2.

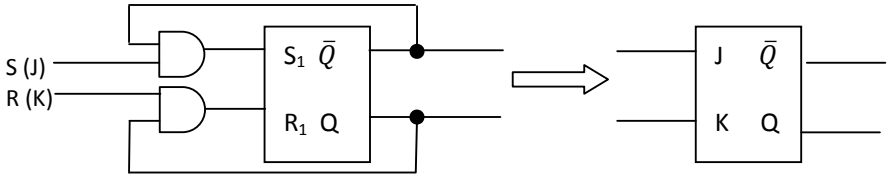


Fig 3.2: Gated-SR-Flip Flop JK-Flip Flop

The Truth Table of Fig 3.2 for the Forbidden states is derived as presented in Table 3.1

Table 3.1: Truth Table of Fig 3.2 when J = k = 1					
S (J)	R (K)	Q_n	$S_1 = S\bar{Q}_n$	$R_1 = RQ_n$	Q_{n+1}
1	1	0	1	0	1
1	1	1	0	1	0
From Truth Table 2.1			From above circuit, Fig 3.2 and equations (3.1) & (3.2)		

When $SRQ_n = 110$. Then,
$S_1 = S\bar{Q}_n = 1 * 1 = 1$
$R_1 = RQ_n = 1 * 0 = 0$
When $SRQ_n = 111$. Then,
$S_1 = S\bar{Q}_n = 1 * 0 = 0$
$R_1 = RQ_n = 1 * 1 = 1$

Comparing the Q_n & Q_{n+1} columns, it shows that the final output, Q_{n+1} of the modified SR-FF TOGGLES when the forbidden states apply as it would do for a JK-FF. That is, when J = K = 1, $Q_n(0) \rightarrow Q_{n+1}(1)$ and $Q_n(1) \rightarrow Q_{n+1}(0)$ showing the toggle-characteristics expected of a conventional JK-Flip Flop.

Option-2 (OR-gate Configuration)

Equations (3.1) & (3.2) are already in AND-gate forms. That is,

$Q_{n+1} = S\bar{Q}_n = S_1$…….. (3.1) $\bar{Q}_{n+1} = RQ_n = R_1$……. (3.2)

From equations (3.1) & (3.2) OR gate equivalent equations can be derived as follows:

$\bar{Q}_{n+1} = \bar{S} + Q_n = S_1$ …... (3.3) $Q_{n+1} = \bar{R} + \bar{Q}_n = R_1$ …….. (3.4)

Hence, the SR-FF can be gated as per these equations as shown in Figure 3.3.

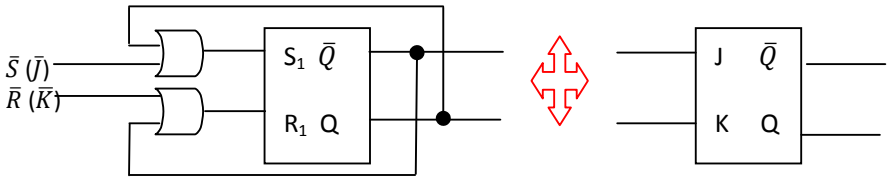


Fig 3.3: Gated-SR-Flip Flop (OR-gate Configuration) JK-Flip Flop

The Truth Table of Fig 3.3 for the forbidden states is derived as presented in Table 3.2

Table 3.2: Truth Table of Fig 3.3 when J = k = 1					
S (J)	R (K)	Q_n	$S_1 = \bar{S} + Q_n$	$R_1 = \bar{R} + \bar{Q}_n$	Q_{n+1}
1	1	0	0	1	0
1	1	1	1	0	1
From Truth Table 2.1			From above circuit, Fig 3.3 and equations (3.3) & (3.4)		

When $SRQ_n = 110$. Then,
$S_1 = \bar{S} + Q_n = 0 + 0 = 0$
$R_1 = \bar{R} + \bar{Q}_n = 0 + 1 = 1$
When $SRQ_n = 111$. Then,
$S_1 = \bar{S} + Q_n = 0 + 1 = 1$
$R_1 = \bar{R} + \bar{Q}_n = 0 + 0 = 0$

Comparing the Q_n & Q_{n+1} columns, it shows that the final output, Q_{n+1} of the modified SR-FF fails to toggle. That is, when J = K = 1, $Q_n(0) \rightarrow Q_{n+1}(0)$ and $Q_n(1) \rightarrow Q_{n+1}(1)$ showing

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

the failure to toggle as expected of a conventional JK-Flip Flop. Hence, this configuration CANNOT be used as a JK-Flip Flop.

Option-3 (NAND-gate Configuration)

Equations (3.1) & (3.2) are already in AND-gate forms. That is,

$Q_{n+1} = S\bar{Q}_n = S_1$…….. (3.1) $\bar{Q}_{n+1} = RQ_n = R_1$……. (3.2)

From equations (3.1) & (3.2) NAND gate equivalent equations can be derived as follows:

$\bar{Q}_{n+1} = \overline{S\bar{Q}_n} = S_1$..... (3.5) $Q_{n+1} = \overline{RQ_n} = R_1$.....(3.6)

Hence, the SR-FF can be gated as per these equations as shown in Figure 3.4.

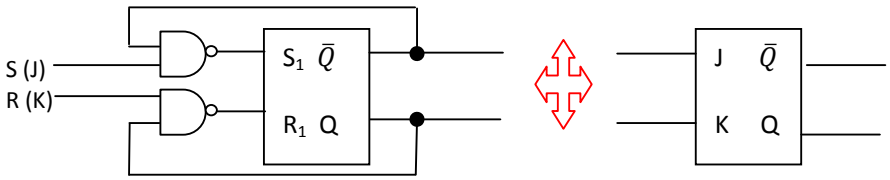


Fig 3.4: Gated-SR-Flip Flop (OR-gate Configuration) ⍰ JK-Flip Flop

The Truth Table of Fig 3.4 for the forbidden states is derived as presented in Table 3.3

Table 3.3: Truth Table of Fig 3.4 when J = k = 1

S (J)	R (K)	Q_n	$S_1 = \overline{S\bar{Q}_n}$	$R_1 = \overline{RQ_n}$	Q_{n+1}
1	1	0	0	1	0
1	1	1	1	0	1
From Truth Table 2.1			From above circuit, Fig 3.4 and equations (3.5) & (3.6)		

When $SRQ_n = 110$. Then,
$S_1 = \overline{S\bar{Q}_n} = \overline{1 * 1} = 0$
$R_1 = \overline{RQ_n} = \overline{1 * 0} = 1$
When $SRQ_n = 111$. Then, $S_1 = \overline{S\bar{Q}_n} = \overline{1 * 0} = 1$
$R_1 = \overline{RQ_n} = \overline{1 * 1} = 0$

Comparing the Q_n & Q_{n+1} columns, it shows that the final output, Q_{n+1} of the modified SR-FF DO NOT CHANGE when the forbidden states apply as against the required toggled action for a JK-FF. That is, when J = K = 1, $Q_n(0) \rightarrow Q_{n+1}(0)$ and $Q_n(1) \rightarrow Q_{n+1}(1)$ showing the failure to toggle as expected of a conventional JK-Flip Flop. Hence, this configuration CANNOT be used as a JK-Flip Flop.

Option-4 (NOR-gate Configuration)

Equations (3.1) & (3.2) are already in AND-gate forms. That is,

$Q_{n+1} = S\bar{Q}_n = S_1$…….. (3.1) $\bar{Q}_{n+1} = RQ_n = R_1$……. (3.2)

From equations (3.1) & (3.2) NOR gate equivalent equations can be derived as follows:

$\bar{Q}_{n+1} = \overline{\bar{S} + Q_n} = S_1$ ….. (3.7) $Q_{n+1} = \overline{\bar{R} + \bar{Q}_n} = R_1$ …… (3.8)

Hence, the SR-FF can be gated as per these equations as shown in Figure 3.5.

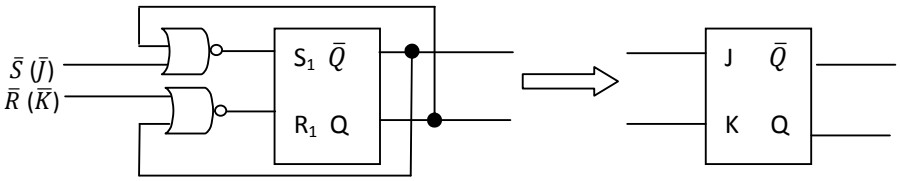


Fig 3.5: Gated-SR-Flip Flop (NOR-gate Configuration) ⍰ JK-Flip Flop

The Truth Table of Fig 3.5 for the forbidden states is derived as presented in Table 3.4

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

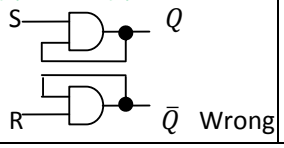
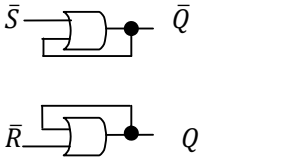
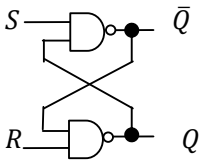
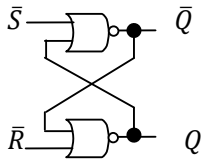
Table 3.4: Truth Table of Fig 3.5 when J = k = 1					
S (J)	R (K)	Q_n	$S_1 = \overline{\bar{S} + Q_n}$	$R_1 = \overline{\bar{R} + \bar{Q}_n}$	Q_{n+1}
1	1	0	1	0	1
1	1	1	0	1	0
From Truth Table 2.1			From above circuit, Fig 3.5 and equations (3.7) & (3.8)		

When SRQ_n = 110. Then,
$S_1 = \overline{\bar{S} + Q_n} = \overline{0 + 0} = 1$
$R_1 = \overline{\bar{R} + \bar{Q}_n} = \overline{0 + 1} = 0$
When SRQ_n = 111. Then,
$S_1 = \overline{\bar{S} + Q_n} = \overline{0 + 1} = 0$
$R_1 = \overline{\bar{R} + \bar{Q}_n} = \overline{0 + 0} = 1$

Comparing the Q_n & Q_{n+1} columns, it shows that the final output, Q_{n+1} of the modified SR-FF TOGGLES when the forbidden states apply as it would do for a JK-FF. That is, when J = K = 1, $Q_n(0) \rightarrow Q_{n+1}(1)$ and $Q_n(1) \rightarrow Q_{n+1}(0)$ showing the toggle-characteristics expected of a conventional JK-Flip Flop.

3. Conversion of an SR-Flip Flop to a JK-Flip Flop Using Table 4.1

Table 4.1: Four Possible Logic Equations from K-Map Considering Only Forbidden States

(b) K-MAP of Q_{n+1}	OR-Gate	NAND-Gate	NOR-Gate
SR: 00 01 11 10 Q_n 0: 0 0 x 1 1: 1 0 x 1 $\bar{Q}_{n+1} = R\bar{Q}_n = S_1 \ldots 4.1$ $Q_{n+1} = SQ_n = R_1 \ldots 4.2$	$\bar{Q}_{n+1} = \bar{S} + \bar{Q}_n = S_1 \ldots (4.3)$ $\bar{Q}_{n+1} = \bar{R} + Q_n = R_1 \ldots (4.4)$	$\bar{Q}_{n+1} = \overline{SQ_n} = S_1 \ldots (4.5)$ $\bar{Q}_{n+1} = \overline{R\bar{Q}_n} = R_1 \ldots (4.6)$	$\bar{Q}_{n+1} = \overline{\bar{S} + \bar{Q}_n} = S_1 \ldots (4.7)$ $\bar{Q}_{n+1} = \overline{\bar{R} + Q_n} = R_1 \ldots (4.8)$
 Wrong	 Right	 Right	 Wrong

This is applicable to NAND-gate SR-Flip Flop configuration. That is, OR-gated SR-Flip Flop and NAND-gated SR-Flip Flop will perform the functions of a conventional JK-Flip Flop.

Since the outputs (Q and $\bar{Q}$) of an SR-FF are complementary to each other, its SR inputs will never be confronted with S = R = 1 if it is triggered by its outputs. Hence, equations (3.1) & (3.2) are used to trigger an SR-FF at the instance when S = R = 1. That is, the SR-FF is triggered during its Forbidden states by its outputs as shown in Fig 4.1.

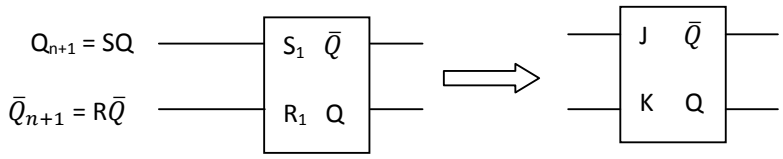


Fig 4.1: Modified SR-Flip Flop ⍰ JK-Flip Flop

Where,
'S' is a new Set-terminal of a modified SR-FF that will perform like a JK-FF (no Forbidden state).
That is, S = J
'R' is a new Reset-terminal of a modified SR-FF that will perform like a JK-FF (no Forbidden state).
That is, R = K
'S_1' is an old Set-terminal of an unmodified SR-FF that that maintains its Forbidden states.
'R_1' is an old Set-terminal of an unmodified SR-FF that that maintains its Forbidden states.

Four different cases/options will be considered using the modified SR-Flip Flop to obtain which of them will behave like a JK-Flip Flop as follows:

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Option-1 (AND-gate Configuration)

Equations (4.1) & (4.2) are already in AND-gate forms. That is,

$Q_{n+1} = SQ = S_1$…….. (4.1) $\bar{Q}_{n+1} = R\bar{Q}_n = R_1$…… (4.2)

Hence, the SR-FF can be gated as per these equations as shown in Figure 3.2.

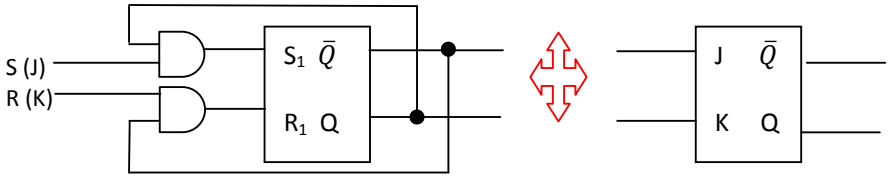


Fig 4.2: Gated-SR-Flip Flop (AND-gate Configuration) ⍰ JK-Flip Flop

The Truth Table of Fig 4.2 for the forbidden states is derived as presented in Table 4.1

Table 4.1: Truth Table of Fig 4.2 when J = k = 1					
S (J)	R (K)	Q_n	$S_1 = SQ$	$R_1 = R\bar{Q}_n$	Q_{n+1}
1	1	0	0	1	0
1	1	1	1	0	1
From Truth Table 2.1			From above circuit, Fig 4.2 and equations (4.1) & (4.2)		

When $SRQ_n = 110$. Then,
$S_1 = S\bar{Q}_n = 1 * 0 = 0$
$R_1 = R\bar{Q}_n = 1 * 1 = 1$
When $SRQ_n = 111$. Then,
$S_1 = S\bar{Q}_n = 1 * 1 = 1$
$R_1 = R\bar{Q}_n = 1 * 0 = 0$

Comparing the Q_n & Q_{n+1} columns, it shows that the final output, Q_{n+1} of the modified SR-FF DO NOT CHANGE when the forbidden states apply as against the required toggled action for a JK-FF.

That is, when J = K = 1, $Q_n(0) \rightarrow Q_{n+1}(0)$ and $Q_n(1) \rightarrow Q_{n+1}(1)$ showing the toggle-characteristics as expected of a conventional JK-Flip Flop are not achieved.

Option-2 (OR-gate Configuration)

Equations (3.1) & (3.2) are already in AND-gate forms. That is,

$Q_{n+1} = SQ = S_1$…….. (4.1) $\bar{Q}_{n+1} = R\bar{Q}_n = R_1$…… (4.2)

From equations (3.1) & (3.2) OR gate equivalent equations can be derived as follows:

$\bar{Q}_{n+1} = \bar{S} + \bar{Q}_n$ ….. (4.3) $Q_{n+1} = \bar{R} + Q_n$ …… (4.4)

Hence, the SR-FF can be gated as per these equations as shown in Figure 4.3.

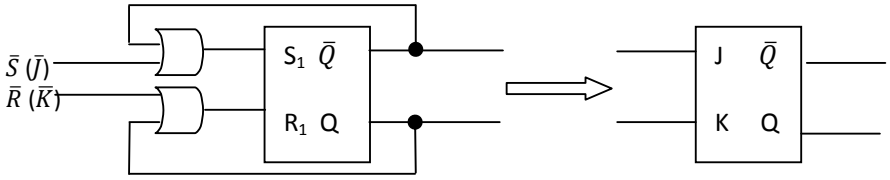


Fig 4.3: Gated-SR-Flip Flop (OR-gate Configuration) ⍰ JK-Flip Flop

The Truth Table of Fig 4.3 for the forbidden states is derived as presented in Table 4.2

Table 4.2: Truth Table of Fig 4.3 when J = k = 1					
S (J)	R (K)	Q_n	$S_1 = \bar{S} + \bar{Q}_n$	$R_1 = \bar{R} + Q_n$	Q_{n+1}
1	1	0	1	0	1
1	1	1	0	1	0
From Truth Table 2.1			From above circuit, Fig 4.3 and equations (4.3) & (4.4)		

When $SRQ_n = 110$. Then,
$S_1 = \bar{S} + \bar{Q}_n = 0 + 1 = 1$
$R_1 = \bar{R} + Q_n = 0 + 0 = 0$
When $SRQ_n = 111$. Then,
$S_1 = \bar{S} + \bar{Q}_n = 0 + 0 = 0$
$R_1 = \bar{R} + Q_n = 0 + 1 = 1$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Comparing the Q_n & Q_{n+1} columns, it shows that the final output, Q_{n+1} of the modified SR-FF fails to toggle. That is, when J = K = 1, $Q_n(0) \rightarrow Q_{n+1}(1)$ and $Q_n(1) \rightarrow Q_{n+1}(0)$ showing the toggle-characteristics as expected of a conventional JK-Flip Flop. Hence, this configuration CAN be used as a JK-Flip Flop.

Option-3 (NAND-gate Configuration)

Equations (4.1) & (4.2) are already in AND-gate forms. That is,

$Q_{n+1} = S\bar{Q}_n = S_1$…….. (3.1) $\bar{Q}_{n+1} = RQ_n = R_1$……. (3.2)

From equations (3.1) & (3.2) NAND gate equivalent equations can be derived as follows:

$\bar{Q}_{n+1} = \overline{SQ_n}$..... (4.5) $Q_{n+1} = \overline{R\bar{Q}_n}$ .....(4.6)

Hence, the SR-FF can be gated as per these equations as shown in Figure 4.4.

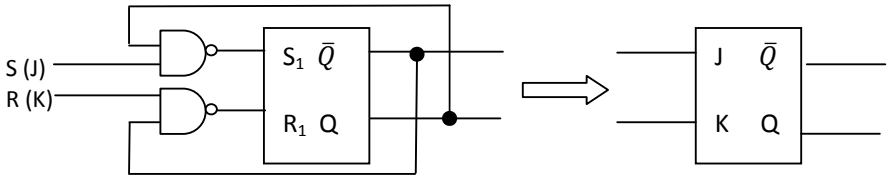


Fig 4.4: Gated-SR-Flip Flop (OR-gate Configuration) ⍰ JK-Flip Flop

The Truth Table of Fig 4.4 for the forbidden states is derived as presented in Table 4.3

Table 4.3: Truth Table of Fig 4.4 when J = k = 1

S (J)	R (K)	Q_n	$S_1 = \overline{S\bar{Q}_n}$	$R_1 = \overline{R\bar{Q}_n}$	Q_{n+1}
1	1	0	1	0	1
1	1	1	0	1	0
From Truth Table 2.1			From above circuit, Fig 4.4 and equations (4.5) & (4.6)		

When SRQ_n = 110. Then,
$S_1 = \overline{SQ_n} = \overline{1 * 0} = 1$
$R_1 = \overline{R\bar{Q}_n} = \overline{1 * 1} = 0$
When SRQ_n = 111. Then, $S_1 = \overline{SQ_n} = \overline{1 * 1} = 0$
$R_1 = \overline{R\bar{Q}_n} = \overline{1 * 0} = 1$

Comparing the Q_n & Q_{n+1} columns, it shows that the final output, Q_{n+1} of the modified SR-FF Toggles when the forbidden states apply as if the Modified SR-Flip Flop were a JK-Flip Flop. That is, when J = K = 1, $Q_n(0) \rightarrow Q_{n+1}(1)$ and $Q_n(1) \rightarrow Q_{n+1}(0)$ as it would have been achieved if conventional JK-Flip Flop were used. Hence, this configuration is suitable to be used as a JK-Flip Flop.

Option-4 (NOR-gate Configuration)

Equations (3.1) & (3.2) are already in AND-gate forms. That is,

$Q_{n+1} = S\bar{Q}_n = S_1$…….. (3.1) $\bar{Q}_{n+1} = RQ_n = R_1$……. (3.2)

From equations (3.1) & (3.2) NOR gate equivalent equations can be derived as follows:

$\bar{Q}_{n+1} = \overline{\bar{S} + Q_n}$ ….. (4.7) $Q_{n+1} = \overline{\bar{R} + \bar{Q}_n}$ …….. (4.8)

Hence, the SR-FF can be gated as per these equations as shown in Figure 4.5.

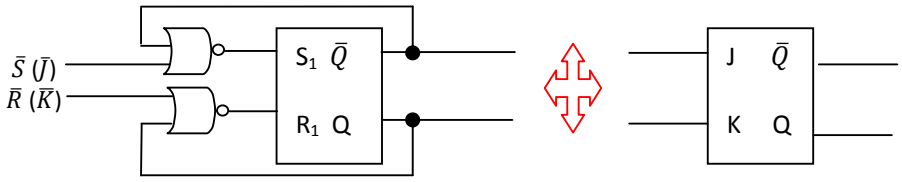


Fig 4.5: Gated-SR-Flip Flop (NOR-gate Configuration) ⍰ JK-Flip Flop

The Truth Table of Fig 4.5 for the forbidden states is derived as presented in Table 4.4

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Table 4.4: Truth Table of Fig 3.5 when J = k = 1					
S (J)	R (K)	Q_n	$S_1 = \overline{\bar{S} + \bar{Q}_n}$	$R_1 = \overline{\bar{R} + Q_n}$	Q_{n+1}
1	1	0	0	1	0
1	1	1	1	0	1
From Truth Table 2.1			From above circuit, Fig 4.5 and equations (4.7) & (4.8)		

When SRQ_n = 110. Then,
$S_1 = \overline{\bar{S} + \bar{Q}_n} = \overline{0+1} = 0$
$R_1 = \overline{\bar{R} + Q_n} = \overline{0+0} = 1$
When SRQ_n = 111. Then,
$S_1 = \overline{\bar{S} + \bar{Q}_n} = \overline{0+0} = 1$
$R_1 = \overline{\bar{R} + Q_n} = \overline{0+1} = 0$

Comparing the Q_n & Q_{n+1} columns, it shows that the final output, Q_{n+1} of the modified SR-FF FAILS TO TOGGLE when the forbidden states apply. That is, when J = K = 1, $Q_n(0) \rightarrow Q_{n+1}(0)$ and $Q_n(1) \rightarrow Q_{n+1}(1)$ showing the toggle-characteristics expected of a conventional JK-Flip Flop has not been achieved.

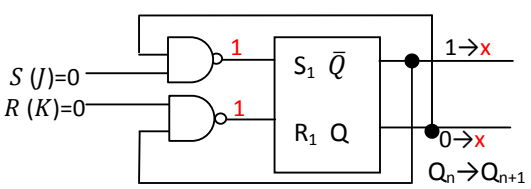
4. Verification of Design

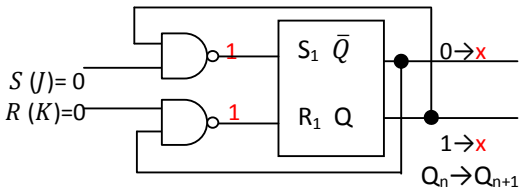
Since the conventional SR-Flip Flop used for the conversion is a NOR-gate Configuration, it is wise to adopt the NOR-gate configuration obtained by the conversion process so that all the gates of the Flip Flop will be of the same type. This will ensure that the propagation time delay and other features will be uniform and this can lead to less error. However, it is necessary to prove the validity of all the designs that appear promising. This is presented in this section.

4.1 NAND-gated-SR-FLIP FLOP CONFIGURATION

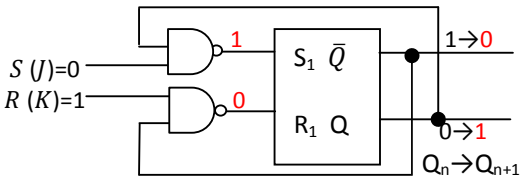
This is presented diagrammatically transitional state by transitional state as follows:

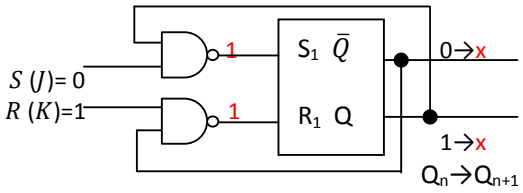
Resting States (0_{10}, 1_{10}): (SRQ_n = 000 & 001)



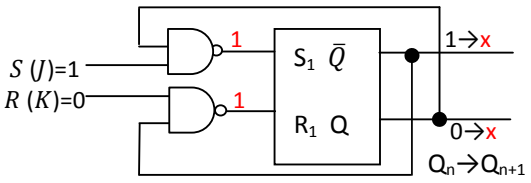


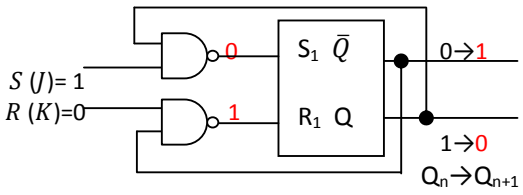
Active States (2_{10}, 3_{10}): (SRQ_n = 010 & 011)



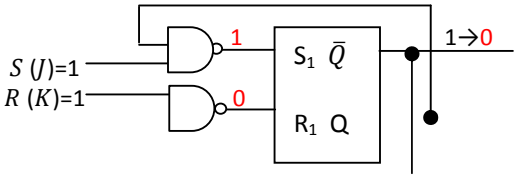


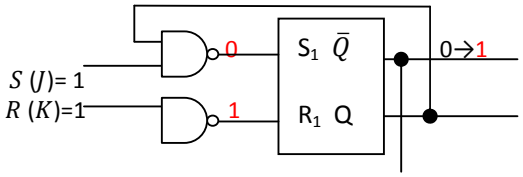
Active States (4_{10}, 5_{10}): (SRQ_n = 100 & 101)



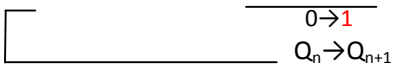


Forbidden States (6_{10}, 7_{10}): (SRQ_n = 110 & 111)





http://sites.google.com/site/ijcsis/
ISSN 1947-5500



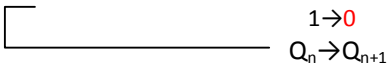
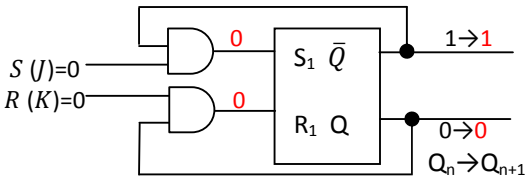
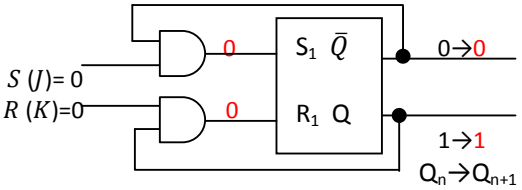


Table 5.1: NAND-gated-SR-Flip Flop											
S/N	CONVENTIONAL SR-FF TRUTH TABLE						NAND-gated-SR-FF TRUTH TABLE				
	S	R	Q_n	Q_{n+1}	REMARKS		S_1	R_1	Q_n	Q_{n+1}	REMARKS
0	0	0	0	0	Resting State		1	1	0	x	Forbidden State
1	0	0	1	1	Resting State		1	1	1	x	Forbidden State
2	0	1	0	0	Active State		1	0	0	1	Active State
3	0	1	1	0	Active State		1	1	1	x	Forbidden State
4	1	0	0	1	Active State		1	1	0	x	Forbidden State
5	1	0	1	1	Active State		0	1	1	0	Active State
6	1	1	0	x	Forbidden State		1	0	0	1	Toggle State
7	1	1	1	x	Forbidden State		0	1	1	0	Toggle State
This configuration satisfies the Toggle states only but failed to satisfy other states. Hence, it cannot be employed.											

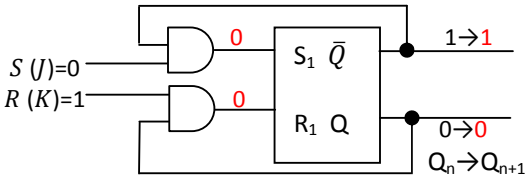
4.2 AND-gated-SR-FLIP FLOP CONFIGURATION

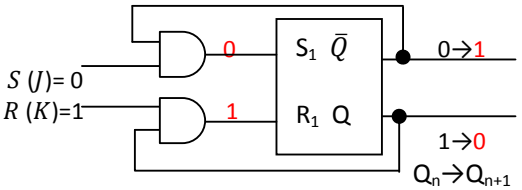
Resting States (0_{10}, 1_{10}): (SRQ_n = 000 & 001)



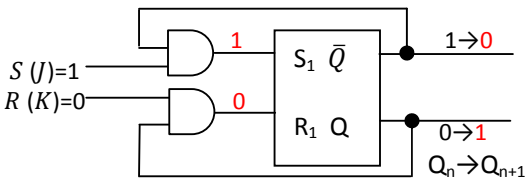


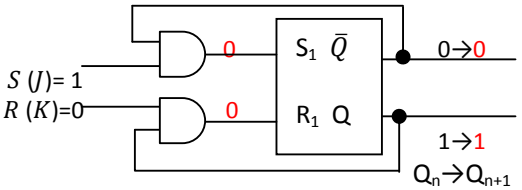
Active States (2_{10}, 3_{10}): (SRQ_n = 010 & 011)



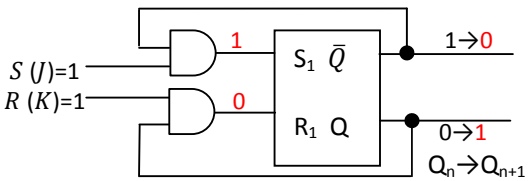


Active States (4_{10}, 5_{10}): (SRQ_n = 100 & 101)





Forbidden States (6_{10}, 7_{10}): (SRQ_n = 110 & 111)



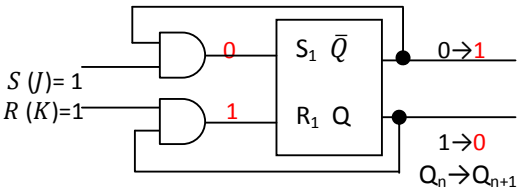


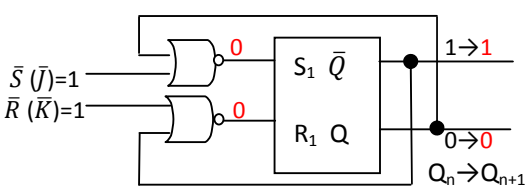
Table 5.2: AND-gated-SR-Flip Flop			
S/N	CONVENTIONAL SR-FF TRUTH TABLE		AND-gated-SR-FF TRUTH TABLE

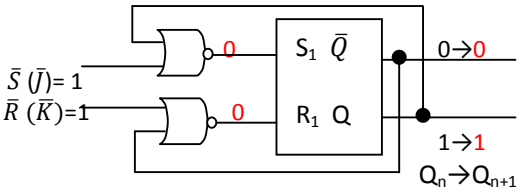
http://sites.google.com/site/ijcsis/
ISSN 1947-5500

	S	R	Q_n	Q_{n+1}	REMARKS		S_1	R_1	Q_n	Q_{n+1}	REMARKS
0	0	0	0	0	Resting State		0	0	0	0	Resting State
1	0	0	1	1	Resting State		0	0	1	1	Resting State
2	0	1	0	0	Active State		0	0	0	0	Active State
3	0	1	1	0	Active State		0	1	1	0	Active State
4	1	0	0	1	Active State		1	0	0	1	Active State
5	1	0	1	1	Active State		0	0	1	1	Active State
6	1	1	0	x	Forbidden State		1	0	0	1	Toggle State
7	1	1	1	x	Forbidden State		0	1	1	0	Toggle State
This configuration satisfies all transitional states including the Toggle states. Hence, it can be employed.											

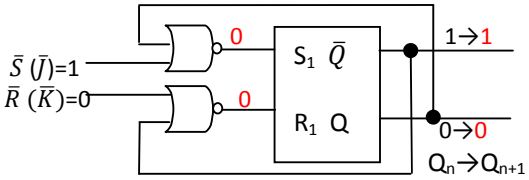
4.3 NOR-gated-SR-FLIP FLOP CONFIGURATION

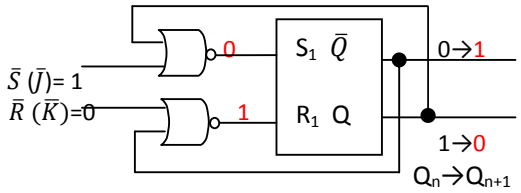
Resting States (0_{10}, 1_{10}): (SRQ_n = 000 & 001)



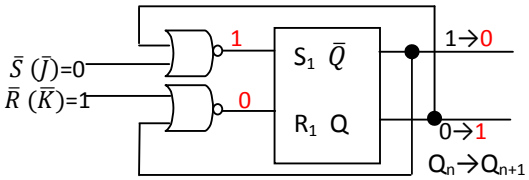


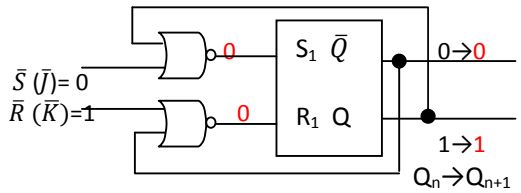
Active States (2_{10}, 3_{10}): (SRQ_n = 010 & 011)



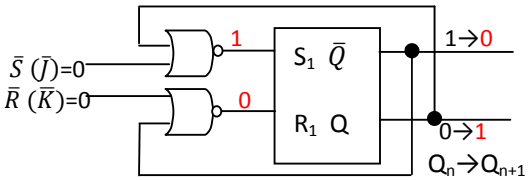


Active States (4_{10}, 5_{10}): (SRQ_n = 100 & 101)





Forbidden States (6_{10}, 7_{10}): (SRQ_n = 110 & 111)



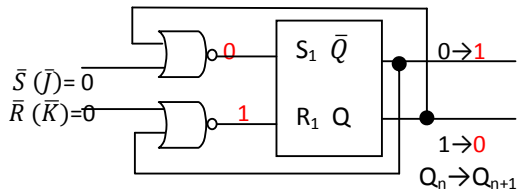


Table 5.3: NOR-gated-SR-Flip Flop

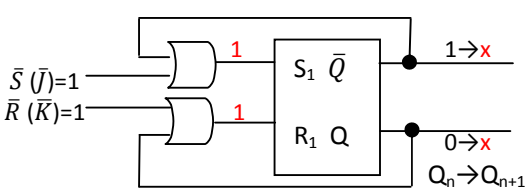
S/N	CONVENTIONAL SR-FF TRUTH TABLE						NOR-gated-SR-FF TRUTH TABLE				
	S	R	Q_n	Q_{n+1}	REMARKS		S_1	R_1	Q_n	Q_{n+1}	REMARKS
0	0	0	0	0	Resting State		0	0	0	0	Resting State
1	0	0	1	1	Resting State		0	0	1	1	Resting State
2	0	1	0	0	Active State		0	0	0	0	Active State

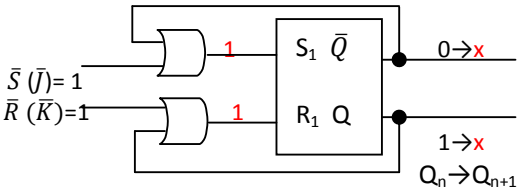
http://sites.google.com/site/ijcsis/
ISSN 1947-5500

3	0	1	1	0	Active State		0	1	1	0	Active State
4	1	0	0	1	Active State		1	0	0	1	Active State
5	1	0	1	1	Active State		0	0	1	1	Active State
6	1	1	0	x	Forbidden State		1	0	0	1	Toggle State
7	1	1	1	x	Forbidden State		0	1	1	0	Toggle State
This configuration satisfies all transitional states including the Toggle states. Hence, it can be employed.											

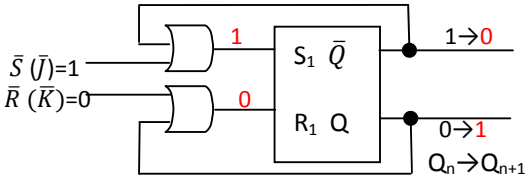
4.4 OR-gated-SR-FLIP FLOP CONFIGURATION

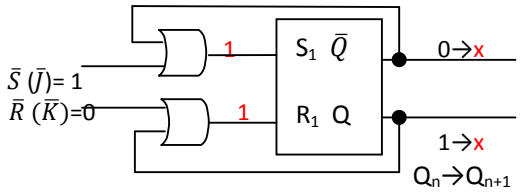
Resting States (0_{10}, 1_{10}): (SRQ_n = 000 & 001)



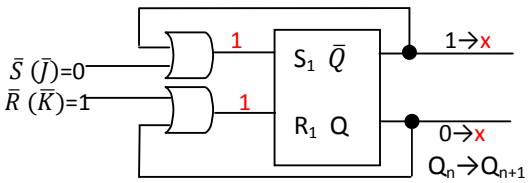


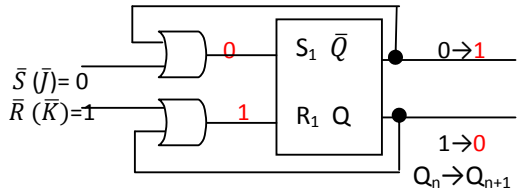
Active States (2_{10}, 3_{10}): (SRQ_n = 010 & 011)



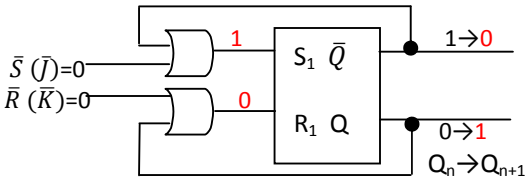


Active States (4_{10}, 5_{10}): (SRQ_n = 100 & 101)





Forbidden States (6_{10}, 7_{10}): (SRQ_n = 110 & 111)



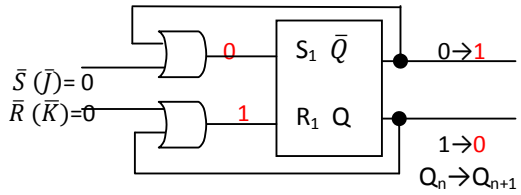


Table 5.4: OR-gated-SR-Flip Flop

S/N	CONVENTIONAL SR-FF TRUTH TABLE						NOR-gated-SR-FF TRUTH TABLE				
	S	R	Q_n	Q_{n+1}	REMARKS		S_1	R_1	Q_n	Q_{n+1}	REMARKS
0	0	0	0	0	Resting State		1	1	0	x	Forbidden State
1	0	0	1	1	Resting State		1	1	1	x	Forbidden State
2	0	1	0	0	Active State		1	0	0	1	Active State
3	0	1	1	0	Active State		1	1	1	x	Forbidden State
4	1	0	0	1	Active State		1	1	0	x	Forbidden State
5	1	0	1	1	Active State		0	1	1	0	Active State

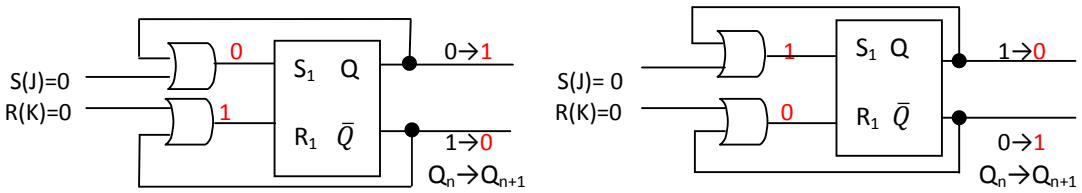
http://sites.google.com/site/ijcsis/
ISSN 1947-5500

6	1	1	0	x	Forbidden State		1	0	0	1	Toggle State
7	1	1	1	x	Forbidden State		0	1	1	0	Toggle State

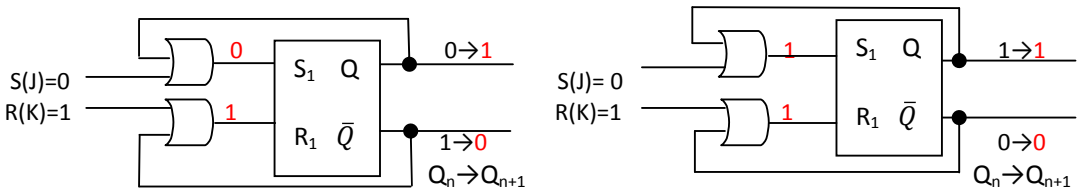
This configuration satisfies Toggle states but fails to satisfy all other transitional states. Hence, it cannot be employed.

4.5 OR-gated-SR-FLIP FLOP CONFIGURATION Using NAND-gate SR-Flip Flop

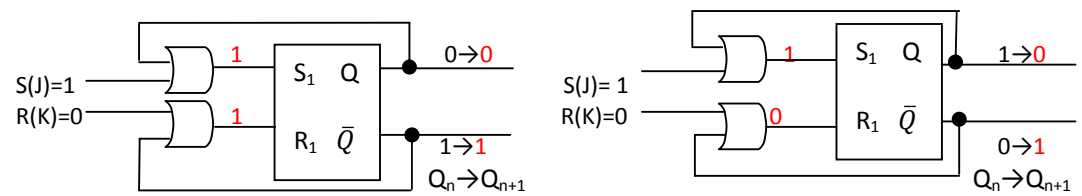
Resting States (0_{10}, 1_{10}): (SRQ_n = 000 & 001)



Active States (2_{10}, 3_{10}): (SRQ_n = 010 & 011)



Active States (4_{10}, 5_{10}): (SRQ_n = 100 & 101)



Forbidden States (6_{10}, 7_{10}): (SRQ_n = 110 & 111)

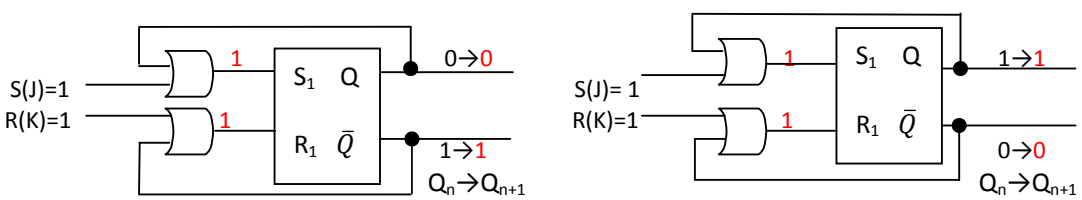


Table 5.5: OR-gated-SR-Flip Flop Using NAND-gate SR-Flip Flop

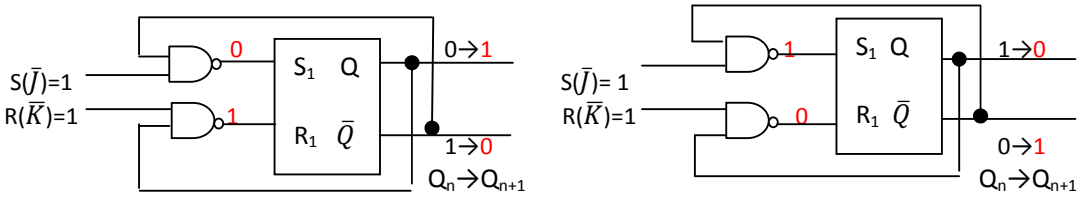
S/N	CONVENTIONAL SR-FF TRUTH TABLE						OR-gated-SR-FF TRUTH TABLE				
	S	R	Q_n	Q_{n+1}	REMARKS		S_1	R_1	Q_n	Q_{n+1}	REMARKS
0	0	0	0	x	Forbidden State		0	1	0	1	Toggle State
1	0	0	1	x	Forbidden State		1	0	1	0	Toggle State
2	0	1	0	0	Active State		0	1	0	1	Active State
3	0	1	1	0	Active State		1	1	1	1	Active State
4	1	0	0	1	Active State		1	1	0	0	Active State
5	1	0	1	1	Active State		1	0	1	0	Active State
6	1	1	0	0	Resting State		1	1	0	0	Resting State
7	1	1	1	1	Resting State		1	1	1	1	Resting State

This configuration satisfies all transitional states including the Toggle states. Hence, it can be employed for a NAND-gate JK-Flip Flop.

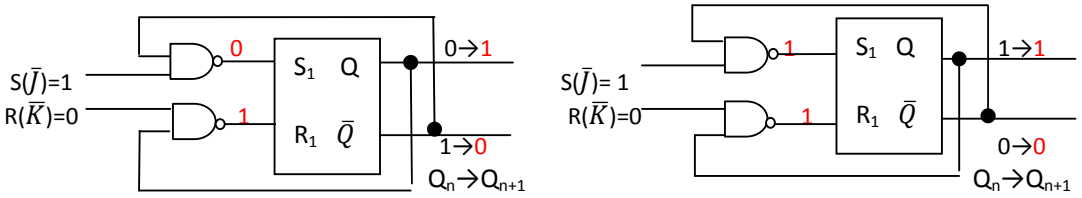
http://sites.google.com/site/ijcsis/
ISSN 1947-5500

4.6 NAND-gated-SR-FLIP FLOP CONFIGURATION Using NAND-gate SR-Flip Flop

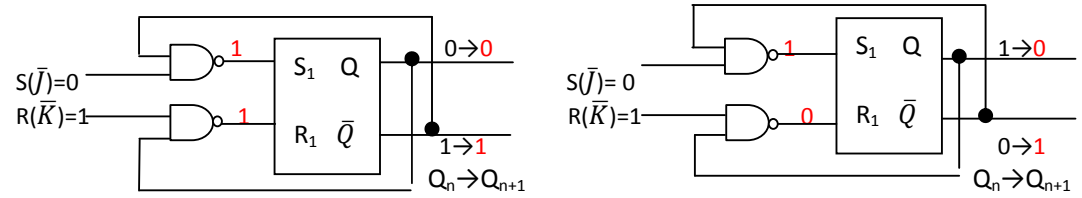
Resting States (0_{10}, 1_{10}): (SRQ_n = 000 & 001)



Active States (2_{10}, 3_{10}): (SRQ_n = 010 & 011)



Active States (4_{10}, 5_{10}): (SRQ_n = 100 & 101)



Forbidden States (6_{10}, 7_{10}): (SRQ_n = 110 & 111)

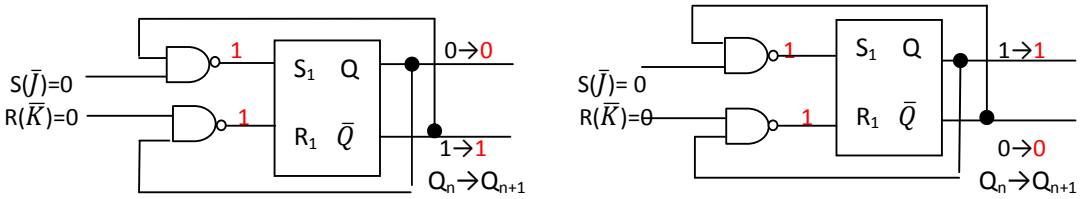


Table 5.6: NAND-gated-SR-Flip Flop Using NAND-gate SR-Flip Flop

S/N	CONVENTIONAL SR-FF TRUTH TABLE						NAND-gated-SR-FF TRUTH TABLE				
	S	R	Q_n	Q_{n+1}	REMARKS		S_1	R_1	Q_n	Q_{n+1}	REMARKS
0	0	0	0	x	Forbidden State		0	1	0	1	Toggle State
1	0	0	1	x	Forbidden State		1	0	1	0	Toggle State
2	0	1	0	0	Active State		0	1	0	1	Active State
3	0	1	1	0	Active State		1	1	1	1	Active State
4	1	0	0	1	Active State		1	1	0	0	Active State
5	1	0	1	1	Active State		1	0	1	0	Active State
6	1	1	0	0	Resting State		1	1	0	0	Resting State
7	1	1	1	1	Resting State		1	1	1	1	Resting State

This configuration satisfies all transitional states including the Toggle states. Hence, it can be employed for a NAND-gate JK-Flip Flop.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

5. Conclusion

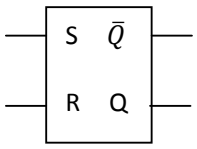
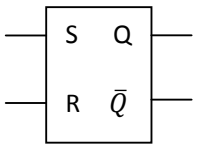
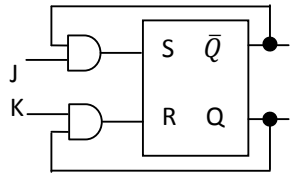
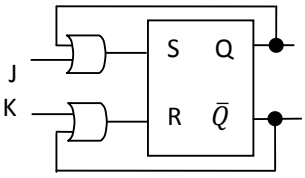
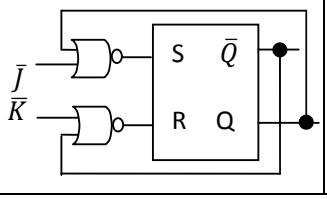
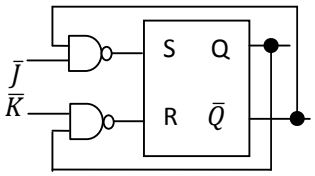
Out of eight possibilities of configuring an SR-Flip Flop to function as a JK-Flip Flop, only four of them appear to be promising that can be used to perform the functions of a JK-Flip Flop. These are as follows:

- AND-gate Configuration with direct output connection. √
- NOR-gate Configurations with cross output connection and complemented J & K inputs.√
- OR-gate Configurations with direct output connection and complemented J & K inputs.
- NAND-gate Configurations with cross output connection.

However, after further analysis, it was discovered that only two configurations can adequately perform the functions of a conventional JK-Flip Flop. These are the AND-gate Configuration and the NOR-gate Configuration.

Note that the analysis presented here is based on the premises of a NOR-gate Configuration of a conventional SR-Flip Flop. The same analysis could be done using a NAND-gate Configuration conventional SR-Flip Flop. If this is done, it could be found out that the OR-gate Configuration and NAND-gate Configuration of the Modified SR-Flip Flop will be adequate for its conversion as proved in section 5.5. Hence, the complete summary is presented in Table 6.1.

Table 6.1: Summary of Conversion of SR-Flip Flop into A JK-Flip Flop

NOR-SR-Flip Flop Configuration	NAND-SR-Flip Flop Configuration	REMARKS
		These are conventional SR-Flip Flops.
		These are conventional SR-Flip Flops converted to function like a JK-Flip Flop.
		These are conventional SR-Flip Flops converted to function like a JK-Flip Flop.

Consequently, downtime of equipment failure due to unavailability of the appropriate spare parts can be reduced by applying this knowledge whenever the situation occurs. By so doing, uninterrupted production due to breakdowns of this digital device can be totally minimised if not avoided.

The analysis here presented also confirms the fundamental postulate that the complete function of an SR-Flip Flop can be fully determined by four appropriate gates and NOT

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

two. Similarly, six gates are required to meet the design of a JK-Flip Flop and NOT four as commonly found in most digital textbooks [2].

References:

[1] Omotosho J. O. & Ogunlere S. O., (2013), "Design Analysis and Circuit Enhancements of SR-Flip flops", International Journal of Engineering Sciences & Research Technology (IJESRT), ISSN: 2277-9655

[2] Omotosho J. O. & Ogunlere S. O., (2013), "Analysis and Design of Different Flip Flops, Extensions of Conventional JK-Flip flops", International Journal of Engineering Sciences & Research Technology (IJESRT), ISSN: 2277-9655

[3] Yngvar B. (2012), "Ultra Low-voltage Differential Static D Flip-Flop for High Speed Digital Applications", Issue 4, Vol. 6, IJCSSP

[4] Omotosho O. J. (2012), "Fundamentals of Digital Systems", Franco-Ola publishers

[5] Phister M. (1958), "Logical Design of Digital Computers", (http://books.google.com/books)

[6] Ranjan J. M, Tripathy and Vijeta (2012), "Comparison of Conditional Internal Activity Techniques for Low Power Consumption and High Performance Flip-Flops", IJCST, Vol. 3, Issue 2

[7] Jyoti, Tripathy M. R. and Vijeta, (2012), "Comparison of Conditional Internal Activity Techniques for Low Power Consumption and High Performance Flip-Flops", International Journal of Computer Science and Telecommunications, ISSN 2047-3338, Volume 3, Issue 2

[8] Mehta K., Arora N. and Singh B. P. (2011) "Low Power Efficient D Flip Flop Circuit", International Symposium on Devices MEMS, Intelligent Systems & Communication (ISDMISC, Proceedings published by International Journal of Computer Applications (IJCA)

[9] Sharma K. G., Sharma T., Singh B. P., and Sharma M. (2011), "Modified SET D-Flip Flop Design for Low-Power VLSI Applications".

[10] Sarica F., (2004), "Level Restoration and Optimization of Current Mode CMOS Multi-Valued Logic Circuits", M.S. Thesis, Bogazici University

Digital Shorthand Based Text Compression

Yogesh *Rathore*
CSE,UIT, RGPV
Bhopal, M.P., India
.

Dr. Rajeev Pandey
CSE,UIT, RGPV
Bhopal, M.P., India
.

Manish k. Ahirwar
CSE,UIT, RGPV
Bhopal, M.P., India
.

***Abstract*—With the growing demand for text transmission and storage as a result of advent of net technology, text compression has gained its own momentum. usually text is coded in yank traditional Code for data Interchange format. Huffman secret writing or the other run length secret writing techniques compresses the plain text[6][11].**

We have planned a brand new technique for plain text compression, that is especially inspired by the ideas of Pitman Shorthand. In these technique we propose a stronger coding strategy, which can provide higher compression ratios and higher security towards all possible ways in which of attacks while transmission. the target of this method is to develop a stronger transformation yielding larger compression and additional security[11].
The basic idea of compression is to transform text in to some intermediate form, which may be compressed with higher efficiency and more secure encoding, that exploits the natural redundancy of the language in creating this transformation.

Keywords-Compression; Encoding; REL; RLL; Huffman; LZ; LZW;Pitman Shorthand;Compression;

I. INTRODUCTION

Data compression is a method of reducing the size of the information to be stored or to be transmitted through a network. Nearly 70-80 % of the Internet users send and receive text-based documents. There is a growing demand for speedy transmission of data, which can be made possible by achieving compression.

There are many techniques already available to reduce the text into compressed format[4]. Most of these techniques use ASCII format (American Standard Code for Information Interchange) that is an 8-bit code. Each character in a text is encoded in a 8-bit format. ASCII is a well-defined set of codes, which is universally accepted.

Text compression techniques have to be context dependent. In Huffman coding method[4], an input text is scanned once from the beginning till the end and the frequency of occurrence of each character is found (histogram). Subsequently, a new coding scheme is followed - frequently appearing characters will have code with less number of bits and least appearing characters are mapped to codes with more number of bits. More text compression methods are Arithmetic coding,Burrows-Wheeler transform,LZW Coding etc[9].

Pitman Shorthand[1] method of documenting is normally practiced by stenographers to take dictation at speaking speed[2][3]. Obviously English or any other language based character set cannot be used to take notes at such speeds. Pitman Shorthand is a proven solution for this requirement. Special graphical sytnbols are used in this method of representing phonetic compositions of the dictated text for certain interval (may be 500 millisecond). This shorthand representation itself is a compressed and encrypted format of the English text. This is the inspiration for us to extend the concept of Pitman Shorthand[l] to compress the plain English text. In this research, a new set of codes is defined and these codes are used instead of graphic symbols. Compression also serves the purpose of encryption.

II. COMPRESSION & DECOMPRESSION

Compression may be a technology by that one or additional files or directory size will be reduced so it\'s straightforward to handle. the target of compression is to scale back the quantity of bits needed to represent information and to decrease the TRM. Compression is achieved through secret writing information and therefore the information is decompressed to its original kind by decryption. Compression will increase the capability of a line by transmittal the compressed file. a standard compressed file that is employed day-today has extensions that finish with .Sit, .Tar, .Zip;

There are two main types of data compression: lossy and lossless.

A. Lossless Compression Techniques

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Lossless compression techniques resurface the initial information from the compressed file with none loss of knowledge[17]. so the knowledge doesn\'t alter throughout the compression and decompression processes. lossless compression techniques square measure accustomed compress pictures, text and medical pictures preserved for jural reasons, laptop viable file then on[9][15].

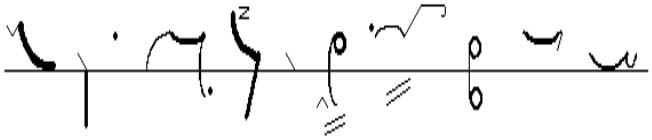
B. Lossy compression techniques

Lossy compression techniques resurface the original message with loss of some information. It is not possible to resurface the original message using the decoding process. The decompression process results an nearly realignment. It may be desirable, when data of some ranges which could not recognized by the human brain can be ignored. Such techniques could be used for multimedia audio, video and images to achieve more compact data compression[7][8].

III. INTRODUCTION TO SHORTHAND METHOD

Shorthand is an abbreviated symbolic writing method that increases speed and brevity of writing as compared to a normal method of writing a language. The process of writing in shorthand is called stenography, from the Greek *stenos* (narrow) and *graphē* or *graphie*. It has also been called brachygraphy, from Greek *brachys* (short) and tachygraphy, from Greek *tachys* (swift, speedy), depending on whether compression or speed of writing is the goal. Fig 1 shows some sentences written in shorthand method for some English statements[1][2][3][16].

Fig 1.Pitman shorthand sentences for some English statements



eye V to D a L-NG-thi V-oiuh-J to S-ow-th aMRKn STS N-the N-you

"I have to do a lengthy voyage to South American cities in the new

'R, to SeeK N-you MahRKtS F R PrawDKt L-eye-n,

year, to seek new markets for our product line."

IV. PROPOSED APPROACH

The basic philosophy of compression is to remodel text in to some intermediate form, which can be compressed with better efficiency and more secure encoding, which exploits the natural redundancy of the language in making this transformation[19].

The frequency of occurrence of each word is found. Subsequently, a new coding scheme is followed - frequently appearing word will have code with less number of special character and least appearing word are mapped to codes with more number of special character combination. Following example show some transformation.

The = ! **said = ***

And = “ **light = :**

God = >)

The algorithm we developing is a three step process consisting:

Step1: Make a Table.

Step2: Encode the input text data.

Step3: Extra Compression by using existing method

Step1: Make a Table

1. Read all words one by one from input files and put in a table.
2. If a word is already within the table increment the quantity of incidence by one, otherwise add it to the table and set the quantity incidence to one.
3. currently kind the table by frequency of occurrences in raining order.
4. begin giving codes victimization the subsequent method:

i). offer the primary 153 words every one permutation of 1 of the ASCII characters.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

ii). currently offer the remaining words every one permutation of 2 of the ASCII characters (in the vary thirty three to sixty four and 128 to 248), taken so as.

If there ar any remaining words offer them every one permutation of 3 of the ASCII characters and finally if needed permutation of 4 characters.

5. produce a brand new table having solely words and their codes. Store this table because the wordbook in a very file.

6. Stop.

Step2: write in code the input text knowledge

1.While computer file isn\'t empty

i. browse the characters from computer file .

ii. If the token is longer than one character, then

a. rummage around for the token within the table

If it's not found,

Write the token as such in to the computer file.

Else

Write corresponding kind word into computer file.

iii. Else

Write corresponding word into computer file.

2. Rename the computer file.

3. Exit

Step3: Extra Compression by using existing method.

we using Gzip for extra compression.

V. PERFORMANCE ANALYSIS

The method is implemented using Java language and the input is tested mainly with three different types of texts - namely running text, (which is normally used in e-mails) addresses and bullet texts. The performance of the algorithm for three different types of text examples is shown in table. A text with rich grammalogues gives highest compression.

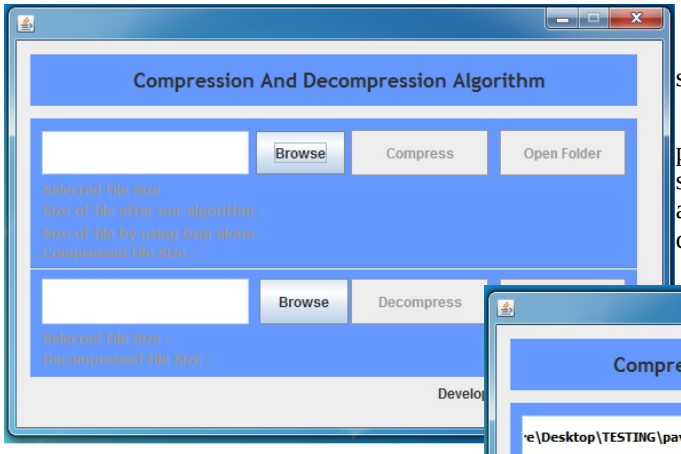


Fig 2. First Windows of software

In this fig first step of our program.Firstly this program show one menu for compression and decompression .Firstly we chose compression optintion.

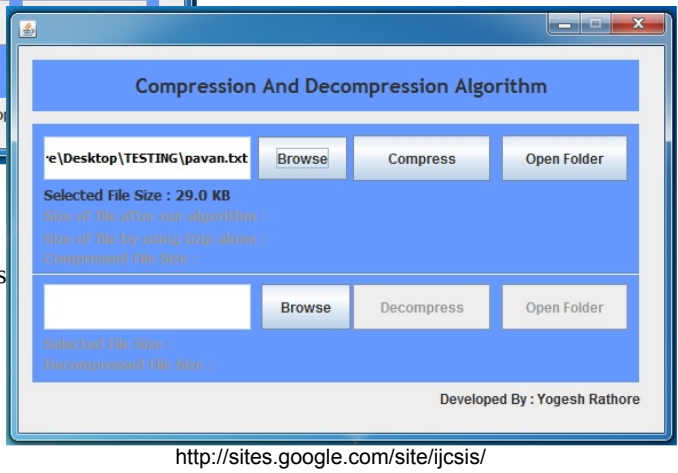


Fig 3. Selecting Compressing File

In this step our program for selecting a file for compression.This windows showes selected file size in back color.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

After selecting compression file now press Compress Button. Now program compressed the file and show compresed file size usin our algoritham and using Gzip algoritham alone. It is show in figure 4.

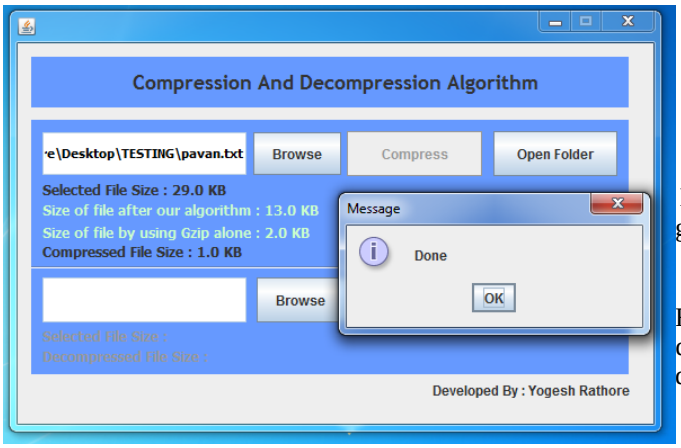


Fig 4 show compreesion done Message.

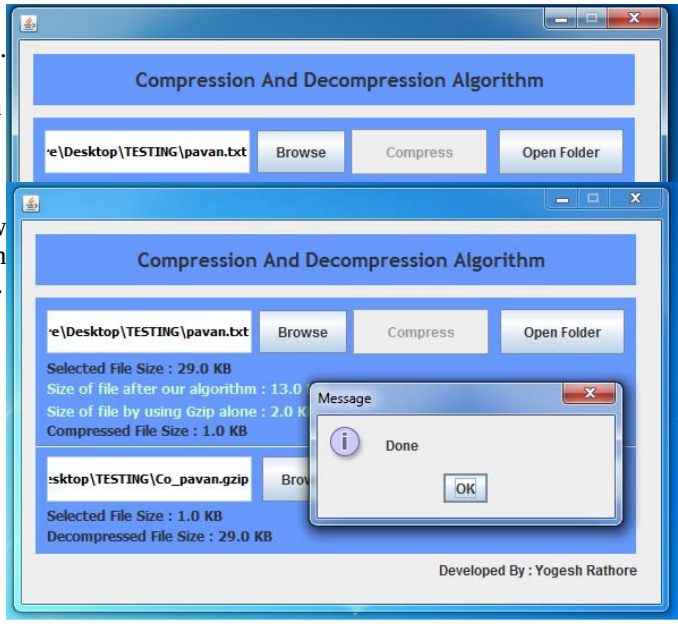
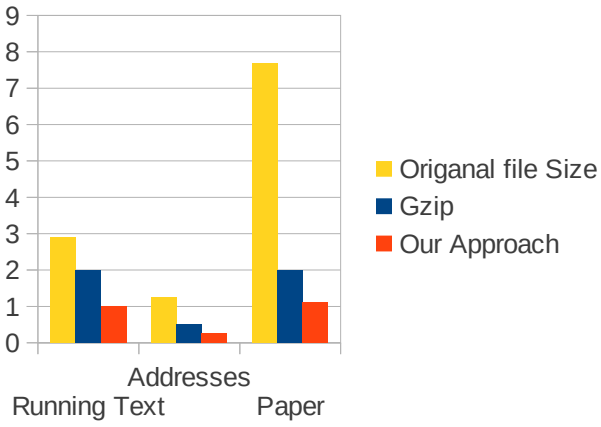


Fig 5. Selecting Decompressing File

Fig 4 show decompreesion done Message.

S. No.	Type Of Text	File size	Average compression in Percentage by Gzip	Average compression in Percentage by Proposed Compression approach
1	Running text	29 kb	93.10%	96.55%
2	Addresses	1.25Kb	52.1%	61.20%
3	Book	7.0 MB	72.28%	85.71%
4	Paper	76.8 kb	72.26%	86.58%



VI. Conclusion

The proposed method based on the concepts of Shorthand Writing Method. It is able to provide compression 15-50% more then to other text compression method. The list of grammalogues used can be changed depending on the situation. However there are certain limitations while coding using this method such as coding .In future work it is worthy to remove its limitation.

References

[1] HemanthaKumar ***G.,*** PhD thesis supervision by Dr. Nagabhushan., On Automation of Text Production from Pitman Shorthand Notes . PhD thesis, University of Mysore, Mysore. 1998.

[2] Isaac Pitman, Shorthand Instructor and Key., Wheeler and Co., 1989

[3] Leedham C G and **A** Nair " Evaluation of dynamic programming algorithms for the recognition of Short forms in itman's shorthand." Journal of Pattern Recognition Letters, Vol 13, 1992

[4] Gonzalvez " Image processing " McGraw-Hill publications 2"d edition

[5] Nagabhushan P and Murali S "Detection of intersection and sequence of stroke segments in Pitman's shorthand document using Hough Transformation", International conference on Cognitive Systems -1999.

[6] Knuth, D. E. 1985. Dynamic Huffman Coding. J. Algorithms 6, 2 (June), 163-180.

[7] Llewellyn, J. A. 1987. Data Compression for a Source with Markov Characteristics. Computer J. 30, 2, 149-156.

[8] Pasco, R. 1976. Source Coding Algorithms for Fast Data Compression.Ph. D. Dissertation, Dept. of Electrical Engineering, Stanford Univ., Stanford, Calif.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[9] Rissanen, J. J. 1983. A Universal Data Compression System. IEEE

[10] Trans. Inform. Theory 29, 5 (Sept.), 656-664.

[11] Tanaka, H. 1987. Data Structure of Huffman Codes and Its Application to Efficient Encoding and Decoding. IEEE Trans. Inform. Theory 33,1 (Jan.), 154-156.

[12] Ziv, J., and Lempel, A. 1977. A Universal Algorithm for Sequential Data Compression. IEEE Trans. Inform. Theory 23, 3 (May), 337-343.

[13] Giancarlo, R., D. Scaturro, and F. Utro. 2009. Textual data compression in computational biology: a synopsis. Bioinformatics 25 (13): 1575-1586.

[14] Burrows M., and Wheeler, D. J. 1994. A Block-Sorting Lossless Data Compression Algorithm. SRC Research Report 124, Digital Systems Research Center.

[15] S. R. Kodifuwakku and U. S. Amarasinge, “Comparison of loosless data compression algorithms for text data”.IJCSE Vol 1 No 4416-225.

[16] en.wikipedia.org/wiki/shothand

[17] RISSANEN, J., AND LANGDON, G. G. 1979. Arithmetic coding. IBM J. Res. Dev. 23, 2 (Mar.), 149-162.

[18] RODEH, M., PRATT, V. R., AND EVEN, S. 1981. Linear algorithm for data compression via string matching. J. ACM 28, 1 (Jan.), 16-24.

[19] Bell, T., Witten, I., Cleary, J., "Modeling for Text Compression", ACM Computing Surveys, Vol. 21, No. 4 (1989).

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

IJCSIS AUTHORS' & REVIEWERS' LIST

Assist Prof (Dr.) M. Emre Celebi, Louisiana State University in Shreveport, USA
Dr. Lam Hong Lee, Universiti Tunku Abdul Rahman, Malaysia
Dr. Shimon K. Modi, Director of Research BSPA Labs, Purdue University, USA
Dr. Jianguo Ding, Norwegian University of Science and Technology (NTNU), Norway
Assoc. Prof. N. Jaisankar, VIT University, Vellore,Tamilnadu, India
Dr. Amogh Kavimandan, The Mathworks Inc., USA
Dr. Ramasamy Mariappan, Vinayaka Missions University, India
Dr. Yong Li, School of Electronic and Information Engineering, Beijing Jiaotong University, P.R. China
Assist. Prof. Sugam Sharma, NIET, India / Iowa State University, USA
Dr. Jorge A. Ruiz-Vanoye, Universidad Autónoma del Estado de Morelos, Mexico
Dr. Neeraj Kumar, SMVD University, Katra (J&K), India
Dr Genge Bela, "Petru Maior" University of Targu Mures, Romania
Dr. Junjie Peng, Shanghai University, P. R. China
Dr. Ilhem LENGLIZ, HANA Group - CRISTAL Laboratory, Tunisia
Prof. Dr. Durgesh Kumar Mishra, Acropolis Institute of Technology and Research, Indore, MP, India
Jorge L. Hernández-Ardieta, University Carlos III of Madrid, Spain
Prof. Dr.C.Suresh Gnana Dhas, Anna University, India
Mrs Li Fang, Nanyang Technological University, Singapore
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Dr. Siddhivinayak Kulkarni, University of Ballarat, Ballarat, Victoria, Australia
Dr. A. Arul Lawrence, Royal College of Engineering & Technology, India
Mr. Wongyos Keardsri, Chulalongkorn University, Bangkok, Thailand
Mr. Somesh Kumar Dewangan, CSVTU Bhilai (C.G.)/ Dimat Raipur, India
Mr. Hayder N. Jasem, University Putra Malaysia, Malaysia
Mr. A.V.Senthil Kumar, C. M. S. College of Science and Commerce, India
Mr. R. S. Karthik, C. M. S. College of Science and Commerce, India
Mr. P. Vasant, University Technology Petronas, Malaysia
Mr. Wong Kok Seng, Soongsil University, Seoul, South Korea
Mr. Praveen Ranjan Srivastava, BITS PILANI, India
Mr. Kong Sang Kelvin, Leong, The Hong Kong Polytechnic University, Hong Kong
Mr. Mohd Nazri Ismail, Universiti Kuala Lumpur, Malaysia
Dr. Rami J. Matarneh, Al-isra Private University, Amman, Jordan
Dr Ojesanmi Olusegun Ayodeji, Ajayi Crowther University, Oyo, Nigeria
Dr. Riktesh Srivastava, Skyline University, UAE
Dr. Oras F. Baker, UCSI University - Kuala Lumpur, Malaysia
Dr. Ahmed S. Ghiduk, Faculty of Science, Beni-Suef University, Egypt
and Department of Computer science, Taif University, Saudi Arabia
Mr. Tirthankar Gayen, IIT Kharagpur, India
Ms. Huei-Ru Tseng, National Chiao Tung University, Taiwan

CALL FOR PAPERS
International Journal of Computer Science and Information Security

IJCSIS 2014
ISSN: 1947-5500
http://sites.google.com/site/ijcsis/

International Journal Computer Science and Information Security, IJCSIS, is the premier scholarly venue in the areas of computer science and security issues. IJCSIS 2011 will provide a high profile, leading edge platform for researchers and engineers alike to publish state-of-the-art research in the respective fields of information technology and communication security. The journal will feature a diverse mixture of publication articles including core and applied computer science related topics.

Authors are solicited to contribute to the special issue by submitting articles that illustrate research results, projects, surveying works and industrial experiences that describe significant advances in the following areas, but are not limited to. Submissions may span a broad range of topics, e.g.:

Track A: Security

Access control, Anonymity, Audit and audit reduction & Authentication and authorization, Applied cryptography, Cryptanalysis, Digital Signatures, Biometric security, Boundary control devices, Certification and accreditation, Cross-layer design for security, Security & Network Management, Data and system integrity, Database security, Defensive information warfare, Denial of service protection, Intrusion Detection, Anti-malware, Distributed systems security, Electronic commerce, E-mail security, Spam, Phishing, E-mail fraud, Virus, worms, Trojan Protection, Grid security, Information hiding and watermarking & Information survivability, Insider threat protection, Integrity
Intellectual property protection, Internet/Intranet Security, Key management and key recovery, Language-based security, Mobile and wireless security, Mobile, Ad Hoc and Sensor Network Security, Monitoring and surveillance, Multimedia security ,Operating system security, Peer-to-peer security, Performance Evaluations of Protocols & Security Application, Privacy and data protection, Product evaluation criteria and compliance, Risk evaluation and security certification, Risk/vulnerability assessment, Security & Network Management, Security Models & protocols, Security threats & countermeasures (DDoS, MiM, Session Hijacking, Replay attack etc,), Trusted computing, Ubiquitous Computing Security, Virtualization security, VoIP security, Web 2.0 security, Submission Procedures, Active Defense Systems, Adaptive Defense Systems, Benchmark, Analysis and Evaluation of Security Systems, Distributed Access Control and Trust Management, Distributed Attack Systems and Mechanisms, Distributed Intrusion Detection/Prevention Systems, Denial-of-Service Attacks and Countermeasures, High Performance Security Systems, Identity Management and Authentication, Implementation, Deployment and Management of Security Systems, Intelligent Defense Systems, Internet and Network Forensics, Large-scale Attacks and Defense, RFID Security and Privacy, Security Architectures in Distributed Network Systems, Security for Critical Infrastructures, Security for P2P systems and Grid Systems, Security in E-Commerce, Security and Privacy in Wireless Networks, Secure Mobile Agents and Mobile Code, Security Protocols, Security Simulation and Tools, Security Theory and Tools, Standards and Assurance Methods, Trusted Computing, Viruses, Worms, and Other Malicious Code, World Wide Web Security, Novel and emerging secure architecture, Study of attack strategies, attack modeling, Case studies and analysis of actual attacks, Continuity of Operations during an attack, Key management, Trust management, Intrusion detection techniques, Intrusion response, alarm management, and correlation analysis, Study of tradeoffs between security and system performance, Intrusion tolerance systems, Secure protocols, Security in wireless networks (e.g. mesh networks, sensor networks, etc.), Cryptography and Secure Communications, Computer Forensics, Recovery and Healing, Security Visualization, Formal Methods in Security, Principles for Designing a Secure Computing System, Autonomic Security, Internet Security, Security in Health Care Systems, Security Solutions Using Reconfigurable Computing, Adaptive and Intelligent Defense Systems, Authentication and Access control, Denial of service attacks and countermeasures, Identity, Route and

Location Anonymity schemes, Intrusion detection and prevention techniques, Cryptography, encryption algorithms and Key management schemes, Secure routing schemes, Secure neighbor discovery and localization, Trust establishment and maintenance, Confidentiality and data integrity, Security architectures, deployments and solutions, Emerging threats to cloud-based services, Security model for new services, Cloud-aware web service security, Information hiding in Cloud Computing, Securing distributed data storage in cloud, Security, privacy and trust in mobile computing systems and applications, **Middleware security & Security features:** middleware software is an asset on
its own and has to be protected, interaction between security-specific and other middleware features, e.g., context-awareness, **Middleware-level security monitoring and measurement:** metrics and mechanisms for quantification and evaluation of security enforced by the middleware, **Security co-design:** trade-off and co-design between application-based and middleware-based security, **Policy-based management:** innovative support for policy-based definition and enforcement of security concerns, **Identification and authentication mechanisms:** Means to capture application specific constraints in defining and enforcing access control rules, **Middleware-oriented security patterns:** identification of patterns for sound, reusable security, **Security in aspect-based middleware:** mechanisms for isolating and enforcing security aspects, **Security in agent-based platforms:** protection for mobile code and platforms, Smart Devices: Biometrics, National ID cards, Embedded Systems Security and TPMs, RFID Systems Security, Smart Card Security, Pervasive Systems: Digital Rights Management (DRM) in pervasive environments, Intrusion Detection and Information Filtering, Localization Systems Security (Tracking of People and Goods), Mobile Commerce Security, Privacy Enhancing Technologies, Security Protocols (for Identification and Authentication, Confidentiality and Privacy, and Integrity), Ubiquitous Networks: Ad Hoc Networks Security, Delay-Tolerant Network Security, Domestic Network Security, Peer-to-Peer Networks Security, Security Issues in Mobile and Ubiquitous Networks, Security of GSM/GPRS/UMTS Systems, Sensor Networks Security, Vehicular Network Security, Wireless Communication Security: Bluetooth, NFC, WiFi, WiMAX, WiMedia, others

This Track will emphasize the design, implementation, management and applications of computer communications, networks and services. Topics of mostly theoretical nature are also welcome, provided there is clear practical potential in applying the results of such work.

Track B: Computer Science

Broadband wireless technologies: LTE, WiMAX, WiRAN, HSDPA, HSUPA, Resource allocation and interference management, Quality of service and scheduling methods, Capacity planning and dimensioning, Cross-layer design and Physical layer based issue, Interworking architecture and interoperability, Relay assisted and cooperative communications, Location and provisioning and mobility management, Call admission and flow/congestion control, Performance optimization, Channel capacity modeling and analysis, Middleware Issues: Event-based, publish/subscribe, and message-oriented middleware, Reconfigurable, adaptable, and reflective middleware approaches, Middleware solutions for reliability, fault tolerance, and quality-of-service, Scalability of middleware, Context-aware middleware, Autonomic and self-managing middleware, Evaluation techniques for middleware solutions, Formal methods and tools for designing, verifying, and evaluating, middleware, Software engineering techniques for middleware, Service oriented middleware, Agent-based middleware, Security middleware, Network Applications: Network-based automation, Cloud applications, Ubiquitous and pervasive applications, Collaborative applications, RFID and sensor network applications, Mobile applications, Smart home applications, Infrastructure monitoring and control applications, Remote health monitoring, GPS and location-based applications, Networked vehicles applications, Alert applications, Embeded Computer System, Advanced Control Systems, and Intelligent Control : Advanced control and measurement, computer and microprocessor-based control, signal processing, estimation and identification techniques, application specific IC's, nonlinear and adaptive control, optimal and robot control, intelligent control, evolutionary computing, and intelligent systems, instrumentation subject to critical conditions, automotive, marine and aero-space control and all other control applications, Intelligent Control System, Wiring/Wireless Sensor, Signal Control System. Sensors, Actuators and Systems Integration : Intelligent sensors and actuators, multisensor fusion, sensor array and multi-channel processing, micro/nano technology, microsensors and microactuators, instrumentation electronics, MEMS and system integration, wireless sensor, Network Sensor, Hybrid

Sensor, Distributed Sensor Networks. Signal and Image Processing : Digital signal processing theory, methods, DSP implementation, speech processing, image and multidimensional signal processing, Image analysis and processing, Image and Multimedia applications, Real-time multimedia signal processing, Computer vision, Emerging signal processing areas, Remote Sensing, Signal processing in education. Industrial Informatics: Industrial applications of neural networks, fuzzy algorithms, Neuro-Fuzzy application, bioInformatics, real-time computer control, real-time information systems, human-machine interfaces, CAD/CAM/CAT/CIM, virtual reality, industrial communications, flexible manufacturing systems, industrial automated process, Data Storage Management, Harddisk control, Supply Chain Management, Logistics applications, Power plant automation, Drives automation. Information Technology, Management of Information System : Management information systems, Information Management, Nursing information management, Information System, Information Technology and their application, Data retrieval, Data Base Management, Decision analysis methods, Information processing, Operations research, E-Business, E-Commerce, E-Government, Computer Business, Security and risk management, Medical imaging, Biotechnology, Bio-Medicine, Computer-based information systems in health care, Changing Access to Patient Information, Healthcare Management Information Technology. Communication/Computer Network, Transportation Application : On-board diagnostics, Active safety systems, Communication systems, Wireless technology, Communication application, Navigation and Guidance, Vision-based applications, Speech interface, Sensor fusion, Networking theory and technologies, Transportation information, Autonomous vehicle, Vehicle application of affective computing, Advance Computing technology and their application : Broadband and intelligent networks, Data Mining, Data fusion, Computational intelligence, Information and data security, Information indexing and retrieval, Information processing, Information systems and applications, Internet applications and performances, Knowledge based systems, Knowledge management, Software Engineering, Decision making, Mobile networks and services, Network management and services, Neural Network, Fuzzy logics, Neuro-Fuzzy, Expert approaches, Innovation Technology and Management : Innovation and product development, Emerging advances in business and its applications, Creativity in Internet management and retailing, B2B and B2C management, Electronic transceiver device for Retail Marketing Industries, Facilities planning and management, Innovative pervasive computing applications, Programming paradigms for pervasive systems, Software evolution and maintenance in pervasive systems, Middleware services and agent technologies, Adaptive, autonomic and context-aware computing, Mobile/Wireless computing systems and services in pervasive computing, Energy-efficient and green pervasive computing, Communication architectures for pervasive computing, Ad hoc networks for pervasive communications, Pervasive opportunistic communications and applications, Enabling technologies for pervasive systems (e.g., wireless BAN, PAN), Positioning and tracking technologies, Sensors and RFID in pervasive systems, Multimodal sensing and context for pervasive applications, Pervasive sensing, perception and semantic interpretation, Smart devices and intelligent environments, Trust, security and privacy issues in pervasive systems, User interfaces and interaction models, Virtual immersive communications, Wearable computers, Standards and interfaces for pervasive computing environments, Social and economic models for pervasive systems, Active and Programmable Networks, Ad Hoc & Sensor Network, Congestion and/or Flow Control, Content Distribution, Grid Networking, High-speed Network Architectures, Internet Services and Applications, Optical Networks, Mobile and Wireless Networks, Network Modeling and Simulation, Multicast, Multimedia Communications, Network Control and Management, Network Protocols, Network Performance, Network Measurement, Peer to Peer and Overlay Networks, Quality of Service and Quality of Experience, Ubiquitous Networks, Crosscutting Themes – Internet Technologies, Infrastructure, Services and Applications; Open Source Tools, Open Models and Architectures; Security, Privacy and Trust; Navigation Systems, Location Based Services; Social Networks and Online Communities; ICT Convergence, Digital Economy and Digital Divide, Neural Networks, Pattern Recognition, Computer Vision, Advanced Computing Architectures and New Programming Models, Visualization and Virtual Reality as Applied to Computational Science, Computer Architecture and Embedded Systems, Technology in Education, Theoretical Computer Science, Computing Ethics, Computing Practices & Applications

Authors are invited to submit papers through e-mail ijcsiseditor@gmail.com. Submissions must be original and should not have been published previously or be under consideration for publication while being evaluated by IJCSIS. Before submission authors should carefully read over the journal's Author Guidelines, which are located at http://sites.google.com/site/ijcsis/authors-notes .

ISSN 1947 5500
http://sites.google.com/site/ijcsis/